THE UNIVERSITY *of* EDINBURGH

Wave radiation in simple geophysical models

Stuart Murray

Doctor of Philosophy
University of Edinburgh
2012

Abstract

Wave radiation is an important process in many geophysical flows. In particular, it is by wave radiation that flows may adjust to a state for which the dynamics is slow. Such a state is described as "balanced", meaning there is an approximate balance between the Coriolis force and horizontal pressure gradients, and between buoyancy and vertical pressure gradients. In this thesis, wave radiation processes relevant to these enormously complex flows are studied through the use of some highly simplified models, and a parallel aim is to develop accurate numerical techniques for doing so.

This thesis is divided into three main parts.

1. We consider accurate numerical boundary conditions for various equations which support wave radiation to infinity. Particular attention is given to discretely non-reflecting boundary conditions, which are derived directly from a discretised scheme. Such a boundary condition is studied in the case of the 1-d Klein-Gordon equation. The limitations concerning the practical implementation of this scheme are explored and some possible improvements are suggested. A stability analysis is developed which yields a simple stability criterion that is useful when tuning the boundary condition. The practical use of higher-order boundary conditions for the 2-d shallow water equations is also explored; the accuracy of such a method is assessed when combined with a particular interior scheme, and an analysis based on matrix pseudospectra reveals something of the stability of such a method.

2. Large-scale atmospheric and oceanic flows are examples of systems with a wide timescale separation, determined by a small parameter. In addition they both undergo constant random forcing. The five component Lorenz-Krishnamurthy system is a system with a timescale separation controlled by a small parameter, and we employ it as a model of the forced ocean by further adding a random forcing of the slow variables, and introduce wave radiation to infinity by the addition of a dispersive PDE. The dynamics are reduced by deriving balance relations, and numerical experiments are used to assess the effects of energy radiation by fast waves.

3. We study quasimodes, which demonstrate the existence of associated Landau poles of a system. In this thesis, we consider a simple model of wave radiation that exhibits quasimodes, that allows us to derive some explicit analytical results, as opposed to physically realistic geophysical fluid systems for which such results are often unavailable, necessitating recourse to numerical techniques. The growth rates obtained for this system, which is an extension of one considered by Lamb, are confirmed using numerical experiments.

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Stuart Murray*)

Acknowledgements

First and foremost, I would like to thank my supervisor Professor Jacques Vanneste, for his guidance, encouragement and seemingly unending patience over the course of my studies at Edinburgh; I am extremely grateful. Throughout my PhD I have had many helpful discussions with a number of people: in particular I would like to thank John Chapman for supplying me with several extremely useful references on Landau poles, as well as Paul Williams and Axel Modave for kindly answering my questions by email. The computing staff in the School of Mathematics have been most helpful in dealing with my many technical issues over the past four years.

I was fortunate to receive a Laura Wisewell travel scholarship which funded my trip to the Courant Institute at New York University during March 2010. This is in addition to conference travel funding received from the Society for Industrial and Applied Mathematics, the London Mathematical Society, the Oxford Centre for Nonlinear PDE, the Wave Flow Interactions Network and the School of Mathematics at Edinburgh. My PhD studies were made possible by funding from the Natural Environment Research Council.

Thanks must go to my fellow postgraduate students (the PGs) in the School of Mathematics, for the lunchtime chats, crossword sessions, sailing adventures, sophisticated parties and everything else that has provided an often much-needed break from mathematics. I have been very fortunate to share an office with Andrew Thompson who has a taste in humour similar to my own (perhaps partly explaining why so many other office mates were dissuaded from continuing their stay with us!), and whose friendship and quiet encouragement has made days spent at the desk all the more enjoyable.

This thesis has been made possible only by the encouragement and support of many kind people. Bruce Duncan has been an invaluable source of compiler advice, coffee and conversation over the past four years. Alistair Mills kindly read a large chunk of this thesis extremely closely, correcting numerous errors and removing any number of barbarisms and solecisms along the way, and has been the best of friends when it was most needed. I am fortunate to have a very supportive family behind me, and in particular must thank my sister and parents for everything they have done for me. Finally, my PhD like everything I do, was made possible by my wife Imogen, who I love dearly and to whom I am so very grateful. She knows what she did.

Contents

Chapter 1

Introduction

1.1 Oceans, the atmosphere and waves

Waves are ubiquitous in geophysical fluids. When thinking of the ocean, we immediately bring to mind those most familiar of waves — *surface gravity waves* — that are driven by surface winds, which we see reaching our shores and giving up their energy through breaking. In the ocean interior we find internal ocean waves: *inertial oscillations* due to the rotation of the planet have a timescale of one day, whereas other internal waves have much higher frequencies. Now add some simple features to our picture of the earth: by introducing coastlines that act as ocean boundaries, the model exhibits *Kelvin waves* which propagate along coastlines, moving anticlockwise in the Northern Hemisphere, and clockwise in the Southern Hemisphere. By improving the model to allow for the fact that the planet is rotating locally faster at the poles than the equator, waves known as *Rossby waves* can be shown to exist. These waves manifest themselves as planetary scale eastward-propagating waves, observed through anomalies in sea surface height.

In the atmosphere we find waves which are perhaps less familiar to most people. Analogous to internal waves in the ocean are *gravity waves*, occasionally made visible by clouds, such as orographic clouds associated with mountain lee waves. Like the ocean, the atmosphere also exhibits inertial oscillations and acoustic waves, with the best example of the latter perhaps being the sound produced by a tornado. A further type of wave — the atmospheric Rossby wave — plays a crucial role in the general circulation, specifically by breaking in the upper atmosphere. Rossby waves propagate upwards from the troposphere (the lowest level of the atmosphere) to the stratosphere, where they break, and in doing so transfer westward momentum.

Rossby wave breaking is perhaps the most pertinent example, because it is an instance of wave radiation being involved as the important catalyst in some dynamical process. In this thesis, wave radiation and its influence on the dynamics is considered for some simple models, all of which are geophysically motivated, with varying degrees of realism. In parallel, we consider some numerical techniques for studying such problems, paying particular attention to accurately simulating wave radiation. Probably the most important of the models involved is the *shallow water model*[64], which we now introduce.

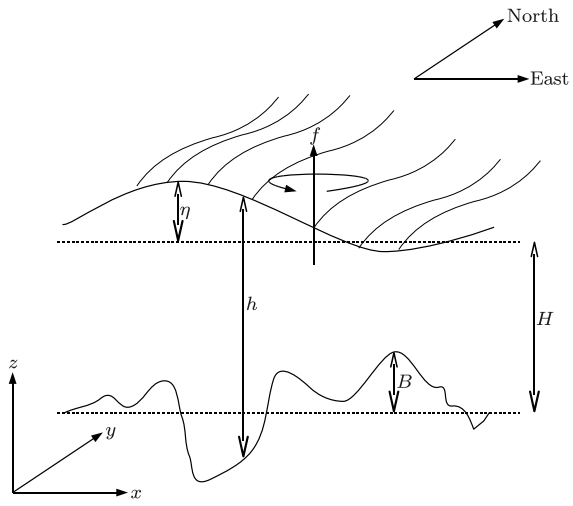


Figure 1.1: A cross section of our fluid layer, showing the various depth measurements used in the rotating shallow water equations, and the coordinate system.

1.2 The rotating shallow water equations

We may characterise the local rate of rotation of the planet using the *Coriolis parameter* denoted by f, which is defined

$$f = 2\Omega \sin(\phi), \tag{1.1}$$

where $\Omega = 2\pi/\text{day}$ and ϕ is latitude. When considering the shallow water equations over spatial scales small enough that variation in f is unlikely to significantly affect the dynamics, it is common to replace the nonlinear expression by a linear or even constant approximation, depending on the spatial scale of interest. Consider a fluid layer rotating with constant Coriolis parameter f that does not depend on position (the so called f-plane approximation). The shallow water approximation involves taking an inviscid fluid of uniform density ρ_0, and setting the horizontal velocity fields u and v to be independent of the depth coordinate z. The associated horizontal velocity vector is written as $\mathbf{u}_\text{h}$, and the following notation for height quantities will be used: H is the average height or rest height of the fluid layer, $B(x, y, t)$ describes the bathymetry, and $\eta(x, y, t)$ is the surface anomaly height. Hence, the total depth of the fluid layer at a particular point in time and space is given by

$$h(x, y, t) = H + \eta(x, y, t) - B(x, y, t). \tag{1.2}$$

See figure 1.1. In all cases that we will consider we take $B(x, y, t) = 0$. Now, consider how the velocity field varies with z: we have already imposed $\partial_z u = \partial_z v = 0$, and hence w is linear in z. We may then find the values of w at the surface and bottom of the fluid at any point, by imposing boundary conditions:

$$w(x, y, t)|_\text{surface} = \frac{D(H + \eta)}{Dt} = \frac{D\eta}{Dt}, \tag{1.3}$$

$$w(x, y, t)|_\text{bottom} = \frac{DB}{Dt} = \mathbf{u} \cdot \nabla B, \tag{1.4}$$

where D is the material derivative. The topography is fixed in time — a very reasonable assumption given the relevant timescales of geophysical fluids. If we now combine the linear profile of w and these boundary conditions, it is clear that

$$w = \left(\frac{z-B}{h}\right)\frac{D\eta}{Dt} + \left(\frac{H+B-z}{h}\right)\nabla\cdot\mathbf{u}_{\mathrm{h}}, \tag{1.5}$$

and hence

$$\frac{\partial w}{\partial z} = \frac{1}{h}\frac{Dh}{Dt}. \tag{1.6}$$

Now from incompressible continuity, we have

$$\frac{\partial w}{\partial z} = -\nabla\cdot\mathbf{u}_{\mathrm{h}}, \tag{1.7}$$

and putting this together with the previous result yields

$$\frac{Dh}{Dt} + h\mathbf{u}\cdot\nabla = 0. \tag{1.8}$$

There are no vertical accelerations, so the z component of the Navier-Stokes equations simply reduces to the hydrostatic balance relation:

$$\frac{\partial p}{\partial z} = -\rho g, \tag{1.9}$$

where g is the acceleration due to gravity. Under our assumption of incompressibility, this integrates to

$$p(z) = p_0 - \rho_0 g z. \tag{1.10}$$

Ignoring any such additional pressures, we find that $p|_{z=\eta} = 0$.

The horizontal pressure gradient is independent of depth, and so for the gradient reduces to $\nabla = (\partial_x, \partial_y)$. From (1.10) and the previous pressure condition, we have

$$p(x,y,z) = \rho_0 g(\eta(x,y)) - z, \tag{1.11}$$

and hence the momentum part of the Navier-Stokes equation reduces to

$$\frac{D\mathbf{u}}{Dt} = -\frac{1}{\rho_0}\nabla p = -g\nabla\eta. \tag{1.12}$$

To account for planetary rotation, the momentum equation is altered by the addition of the term $f(\hat{\mathbf{z}}\times\mathbf{u})$. The rotating shallow water equations are then as follows.

The rotating shallow water equations on the f-plane:

$$\frac{\partial h}{\partial t} + \nabla\cdot(h\mathbf{u}) = 0, \tag{1.13}$$

$$\frac{D\mathbf{u}}{Dt} + f(\hat{\mathbf{z}}\times\mathbf{u}) = -g\nabla\eta. \tag{1.14}$$

1.3 Motivation

1.3.1 Solving problems on unbounded domains

The fluid models studied in this thesis all have wave radiation to infinity as a common feature, and hence we require practical numerical methods for solving such unbounded problems given finite computing resources. Part of this thesis therefore considers *non-reflecting boundary conditions*. A non-reflecting boundary condition can be described as one which allows waves to leave the domain but not enter it. However, such a description is neither quite as simple or precise as it would first seem, and as a result there are many methods for deriving such conditions, leading to conditions with a variety of behaviours. Tsynkov [75] divides them into two broad categories: those that are globally exact, and those that are locally exact. In the former, boundary conditions are accurate, but generally computationally expensive, while the latter includes methods that are relatively cheap and geometrically simple, though perhaps not as accurate as may be desired, leading to the problem of finding the right balance between accuracy and practicability. In this thesis, we study methods from each group. Of course, there are other options when simulating problems on an infinite domain; one may use wave absorbing layers (so called "sponge layers"), which are simply regions of artificial damping [47], or else impose periodic boundary conditions. However, for the models which we will study, the former is not as accurate as we would like, and the latter does not allow waves to radiate away from the region of interest.

For the cases we study, it is possible to derive an artificial boundary condition that is perfectly non-reflecting for the continuous equation that is being solved. However, when seeking accurate boundary conditions we must consider the effect of *artificial dispersion* introduced by the process of discretisation, as illustrated by figure 1.2 which shows the dispersion relations obtained for both the continuous and discretised form of Klein-Gordon equation (a central model in this thesis). One motivation is the development of boundary conditions which are matched to the dispersion relation of the numerical scheme, i.e. boundary conditions that perform well regardless of numerical resolution. We pay particular attention to the practical implementation of all the boundary conditions developed; in particular, we study the effects on accuracy of certain practical alterations to NRBCs (e.g. truncation) and the corresponding interior scheme (e.g. filtering), which have been overlooked before, and propose some remedies.

1.3.2 Wave radiation

In this work we concentrate on two important mechanisms that are due to wave radiation: the maintenance of balance in a geophysical model, and *Landau damping*, a peculiar type of fluid damping. In many fluid scenarios, the presence of wave radiation leads to a competition between instability and Landau damping. Examples of competition between such mechanisms can be found in work by Schecter & Montgomery [69] on mesoscale cyclones in shallow water on the f-plane, and Balmforth et al [4], where Landau damping is counteracted by nonlinear effects that encourage growth. Particularly interesting are the cases where the balance is a fine one, with the winning effect decided by some perhaps rather subtle factors. In this thesis, we consider each effect separately.

It is unfortunate that seemingly simple models of fluid flow (such as the shallow water

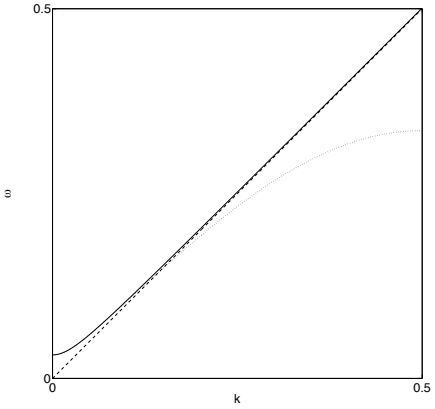


Figure 1.2: The exact dispersion relation for the Klein-Gordon equation (solid curve) and that obtained using a second-order centred-time/centred-space finite-difference scheme (dotted curve), for a particular choice of time and grid step. The wavespeed and dispersion parameter are both unity. Also shown is the dispersionless case, $\omega = k$ (dashed line).

approximation) are often difficult to probe analytically. Rather than immediately resorting to numerical techniques, it is sometimes more helpful to study a model that has similar qualitative features, that is nonetheless easier to deal with. For example, we may introduce toy models perhaps comprising a few ODEs coupled with a 1-d PDE; several novel examples of such models are studied in this work. In addition, model reduction techniques can be used, especially for systems with large timescale separation. The ocean and atmosphere are examples of systems with broadly separated timescales. To see this, note that the shallow water equations possess three branches of solutions: two IGW modes, and a vortical mode. The inertia-gravity waves have frequency $\omega > f$, which compared to the advective frequency $f_{\text{adv}} \sim L/U$ is fast. Such fast motions are relatively weak, with the result that it is possible to represent the dynamics using a *balanced model*, which effectively filters the fast IGW motion from the solution. This corresponds to defining a *slow manifold* in the state space [52, 55, 84], which describes the coupling between the fast and slow variables. There is however a limit to the accuracy of the reduction that can be performed in this way, and there will always be some fast motion that develops no matter how "close" to balance the initial data lies. For this reason, the slow manifold is sometimes described as "fuzzy" or given the name *quasimanifold*. In this thesis, it will be useful to reduce one of the simple geophysical models to a balanced model, given the clear timescale separation. We may use such a balanced model to investigate *whether wave radiation is a mechanism by which a balanced state may be maintained* and the dynamics kept slow in the interior. We also use the model to probe some questions surrounding the way that quantities such as the wave energy flux are related to the slow variables.

1.3.3 Quasimodes

In a classic paper, Briggs et al [6] considered a 2-d, inviscid, compact, axisymmetric vortex with a step vorticity profile, $\zeta(r)$. Such a vortex supports Rossby-Kelvin modes which deform it, and propagate around the boundary in the direction of the flow. However, by introducing even the slightest gradient in the vorticity profile, such Rossby-Kelvin modes are no longer

possible solutions. It seems odd that such a subtle alteration of the vorticity field should have such a result, but in fact the effect is not as drastic as it seems, for we still observe mode-like disturbances, which decay — at least for a short time, and only when examining coarse grained fields — exponentially. In the latter part of this thesis, we focus on these so called "quasimodes" — **disturbances which behave like true modes in a system which does not support eigenmodes** — which have an elegant mathematical description in the form of *Landau poles*. In the model that we use to study quasimodes, the Landau damping is a result of wave radiation, which is in contrast to all of the fluid examples we give. However, the model still provides us with useful insights.

1.4 Questions and objectives: a thesis outline

In this thesis we study several models which demonstrate the importance of wave radiation and Landau damping in some simple systems which are relevant to certain geophysical models. The aim is to assess how wave radiation allows excess energy to be removed from a system, and also maintains a state of balance. Another objective is to improve the understanding of quasimodes and their appearance in fluid systems. In light of the mathematical complexity of some of the fluid systems of interest, several toy models are used instead. Given the nature of the questions being asked, it is crucially important that these models handle wave radiation to infinity both accurately and efficiently, which naturally means giving due consideration to the boundary conditions employed in any numerical scheme used to simulate wave radiation. Hence, a parallel aim of this thesis is to develop new numerical boundary conditions for a variety of wave radiation problems, and to derive some new results for existing ones.

Chapter 2 lays the foundations for the numerical techniques used throughout this thesis, and develops some novel results for non-reflecting boundary conditions. In §1, 1-dimensional boundary conditions are developed for the Klein-Gordon equation, as this is an equation that will feature throughout, due to the similarity between its dispersion relation and that of the shallow water equations. After considering boundary conditions that are derived from the continuous equations, a discretely non-reflecting boundary condition is derived, in an effort to produce a numerical boundary condition that behaves well with coarse numerical resolution. A boundary condition studied recently [38] is reviewed and some modifications and improvements made, and a useful stability argument is developed. In §2, higher order boundary conditions — originally due to Higdon — are studied for use in 2-d shallow water models. A new efficient implementation is first derived, before we take the new step of combining a recent scheme (due to Givoli and Neta [29]) for choosing the free parameters in the boundary condition, with a grid scheme investigated by Dea[15]. The effectiveness of the Givoli-Neta algorithm is then assessed. The stability of the whole scheme is established numerically for the first time.

Having built up a suite of numerical tools for simulating transparent boundaries, chapter 3 sees them applied to answer some questions about competitive forcing and radiative effects in a toy model. A simple five component model due to Lorenz and Krishnamurthy, subsequently extended by Vanneste, is modified by the addition of stochastic forcing. Such a system serves as a caricature of wind forcing of the ocean, and provides new insight since it features dissipation due to wave radiation, rather than damping. Higher order balance relations are derived for

the model, and these are used in a numerical scheme with an accurate non-reflecting boundary condition to examine wave radiation. One aim is to investigate radiation as a mechanism for maintenance of balance, and also to derive a rationale for relating some basic knowledge of the slow dynamics to the outgoing wave flux in the far field. Some of these results are then investigated for fluids, by using a 2-d nonlinear shallow water model with a smooth random wind stress applied at the surface. Following methods used in work by Ford [24] and Schecter [69], we re-derive the unperturbed fields for a 2-d axisymmetric shallow water vortex with a general PV profile, and then use this as the basic state to undergo forcing.

It is generally the case that for fluid systems exhibiting quasimodes, the analytical procedure for calculating their decay rate is rather involved and cumbersome. Therefore, it is both interesting and useful to study simpler systems where such calculations by hand are more straightforward. In chapter 4, we consider a coupled oscillator-string system, which is an extension of one considered by Lamb [50]. Landau poles (corresponding to Landau damping) are uncovered and analytical expressions for their growth rates are derived. These findings are then compared to results from numerical simulations, which make use of those discretely non-reflecting boundary conditions derived in chapter 2. The model is extended further by addition of a simple damping term, and the associated bifurcation is examined. In chapter 5 we present our conclusions.

Chapter 2

Numerical boundary conditions for wave radiation problems

2.1 The need for non-reflecting boundary conditions

The main thrust of this thesis involves the investigation of wave radiation and its effect on the dynamics in a variety of settings, from simple one-dimensional toy models, to nonlinear two-dimensional fluid models. There is an overarching need to create schemes that accurately simulate the feature common to all of the problems being studied, which is that *waves are radiated away from the domain of interest.* Given that such problems involve waves being radiated freely to infinity, ideal numerical methods allow waves generated in the problem domain to likewise propagate freely — leaving the domain of interest unhindered and unaffected by the boundary.

Let us consider the practical implications of such a requirement. All of the numerical simulations will be carried out on a grid, and any such grid will be finite due to the physical memory restrictions of a computer, and will therefore have far-field boundaries. As any child who has dropped a pebble into a small pond knows, a wall at the boundary will cause waves to be reflected back towards their point of origin. Therefore, one strategy for simulating wave radiation to infinity — and one that gives perfect results — is simply to make the computational domain large enough that no wave ever reaches the boundary. The drawback of such a strategy is that very long simulations require a correspondingly large computational domain. For example, a problem that receives attention later is that of a slowly growing vortex instability, which requires solution over long timescales in order to observe any of the interesting dynamics. The computational domain may need to be very large indeed, if any of the waves it supports are fast and would otherwise reach a boundary quickly. For problems in more than one dimension, we cannot continue to increase the size of the computational domain arbitrarily, without incurring severe and unaffordable costs in terms of computation time.

The main alternative to the first approach is to formulate a *non-reflecting boundary condition* which, for a particular model system, allows disturbances to escape from the interior and not return. As we shall see, translating this simple requirement into a workable boundary condition is highly nontrivial. These boundary conditions are equally referred to in the literature

as *transparent*, *absorbing*, or *artificial* boundary conditions. Such a boundary condition should be designed to produce as little reflection as possible, and we may reserve the terms "perfect" or "exactly" non-reflecting for those that produce a solution in the interior identical to that which would be obtained by solving the problem on a large enough domain to avoid reflections. Finally, although not themselves boundary conditions, *sponge layers* or *wave absorbing layers* are a further technique that may be applied to problems of wave radiation. Sponge layers are regions where spatially varying damping is applied to the solution, so that outgoing waves are damped to an acceptably small amplitude before being dealt with by a simple boundary condition, such as a hard wall condition.

Given the importance of an effective i.e. stable and accurate method of simulating wave radiation for the problems tackled in subsequent chapters, they merit some attention. Therefore this chapter is devoted to the choice and study of several non-reflecting boundary conditions for the model problems to be considered. In §2.2, the many types of NRBC are summarised and the literature surveyed. In §2.3 boundary conditions for a one-dimensional dispersive system are derived and analysed, which will later be applied to toy models for wave radiation and Landau damping; §2.4 concerns so-called higher order boundary conditions for the 2-d linearised shallow water equation, while Appendix F deals with sponge layers used in our 2-d nonlinear shallow water equation solver.

2.2 The hierarchy of non-reflecting boundary conditions

The literature on non-reflecting boundary conditions — henceforth abbreviated to NRBC — is broad and slightly dispersed, due to the fact that there are so many possible approaches to their derivation. An overview of NRBCs is given by Hagstrom [34], and more recently in reviews by Nycander and Döös [62], Givoli [28] and Tsynkov [75]. Other useful references include widely cited works by Keller and Givoli [48], Givoli alone [27], and Lill [53]. Durran [19] provides a discussion of applications of non-reflecting boundary conditions. Here the various types of such boundary conditions are reviewed and given their proper place in the hierarchy of NRBCs.

One common procedure for deriving a numerical scheme with reflecting boundary conditions is to begin with the PDEs describing the system to be simulated, and factoring into expressions for ingoing and outgoing modes. An exact boundary condition may then be derived, which can itself be discretised. For example, it is a trivial matter to write the 1-d wave equation as $[(\partial_t + c\partial_x)(\partial_t - c\partial_x)]u(x,t) = 0$; one then obtains the exact non-reflecting boundary conditions

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial t} = 0; \quad \frac{\partial u}{\partial t} - c\frac{\partial u}{\partial t} = 0, \tag{2.1}$$

for the left- and right-hand boundaries respectively. Ehrhardt [21] considers this approach when applied to hyperbolic systems, and gives examples. Staying with the wave equation, Grote & Keller [31] developed boundary conditions for the wave equation in more than one dimension, which proves considerably more difficult than for the above example. For systems with dispersion, Hagstrom [35] and subsequently Han & Zhang (§2 of [38]) derived NRBCs from the linear Klein-Gordon equation.

So far we have considered boundary conditions derived from the governing equations, and have not considered discretisation of either these equations or the NRBC. However, if we naively discretise both the equations for the interior and the NRBC using a particular finite-difference approximation, then there is no guarantee that the resulting boundary condition will not cause reflections when used together with the interior finite-difference scheme in question. Clearly, this approach produces boundary conditions whose accuracy depends on the time step and grid size. If a simulation aims at achieving maximum accuracy with minimal computation, the discretisation of the boundary conditions should be chosen to have to have the same order of accuracy as the scheme used in the interior. However, in many applications, accuracy is not achievable and the focus of simulations shifts to describing as faithfully as possible the qualitative behaviour of the system at hand. In long-time simulations of complex, chaotic systems, in particular, the sensitive dependence of trajectories on initial conditions and small perturbations (including truncation errors) means that one aims at computing statistical properties, such as an invariant measure, rather than individual trajectories. This difference of focus has implications for the numerical methods, emphasising those which capture properties of the continuous systems that are key to their qualitative behaviour such as energy conservation or symplecticity. For example, when considering the shallow water equations, one may wish to choose a scheme with whose geometry and stencil mean that the dispersion properties of the discretisation match those of the governing continuous equations [71]. This principle is well established for ordinary differential equation models (e.g. in celestial and molecular dynamics [36]), and it clearly applies to partial differential equation models (e.g. in climate modelling or quantum chemistry). What is perhaps less clear in the context of partial differential equations is which properties are essential for the long-time qualitative behaviour of solutions. When wave propagation is involved, it is obvious that characteristics such as number of wave modes and general features of the group velocity are important [72]. In an unbounded domain, radiation of the wave energy to infinity is another key property, and it is therefore important to develop radiation boundary conditions that avoid spurious reflection with an accuracy that is higher than that imposed by the discretisation in the interior. One way of achieving this is to employ *discretely* (or *numerically*) non-reflecting boundary conditions (DNRBCs), that is, boundary conditions derived directly from exact solutions of the discretised equations rather than from the continuous equations. The reflection properties of these boundary conditions are decoupled from the truncation error and can be controlled independently, so that one can ensure negligible reflection even for relatively coarse resolutions in the interior. In this section, we focus on DNRBCs for a simple dispersive system in order to assess their potential for applications where, as just described, minimum reflection is prioritised over interior accuracy.

Tuomela & Vacus [76] developed a general framework for the algebraic derivation of perfect DNRBCs for the wave equation. Dispersion adds its own complications when one attempts to find a suitable DNRBC, and many efforts have been made in this direction. Wilson [87] also studied boundary conditions derived directly from the discrete scheme. He worked with a general hyperbolic system discretised using a Lax-Wendroff type difference scheme and was able to derive boundary conditions that are local in time[space] but require full knowledge of the solution in space[time], in one spatial dimension. Rowley & Colonius [67] went on to develop a general technique for finding DNRBCs local in both space and time for linear hyperbolic systems. Most recently, Han & Zhang (§5 of [38]) derived a perfectly non-reflecting DNRBC

for the Klein-Gordon equation solved using a centred-in-time, centred-in-space, second-order finite-difference scheme.

Perfectly or exactly non-reflecting boundary conditions commonly have a the feature of being non-local in space or time. Typically, either the entire history of the solution at a single grid point is required, or knowledge of the solution is required further and further away from the boundary as the simulation progresses. Given the obvious practical drawbacks in computational overhead associated with this class of DNRBC, one can instead *design* an NRBC somewhat heuristically, perhaps by choosing a particular stencil or basing it on an NRBC for the continuous equations. This more practical class of NRBCs has also received much attention. Although no longer perfect, they are apt in situations where an exact DNRBC is impractical, and one would instead prefer an alternative that is simple to implement whilst also offering good reflecting properties.

Engquist & Majda [23] derived several DNRBCs for the wave equation in one and later two dimensions. They found a perfectly absorbing NRBC that was non-local in space, but subsequently took the more practical approach (necessitated by the limitations of computing machines of the day), of assuming a boundary operator involving the boundary point and its neighbour at three time levels, and "tuning" this local operator to match the discrete dispersion relation of the interior difference scheme as closely as possible, thus minimising reflections. Building on the work of Engquist & Majda, Wagatha [83] also studied NRBCs for the wave equation and shallow water equations. Rowley & Colonius also considered local NRBCs constructed along the lines suggested in [23].

Higdon [39, 40] developed boundary conditions for the multi-dimensional wave equation solved using a centred-space/-time scheme, that can be extended to any order based on compositions of approximations to analytical non-reflecting conditions. These have been shown to be very effective, particularly because their reflection properties improve at higher orders, and they are in common use. This work was subsequently extended to deal with dispersive systems [42, 43]. Higdon boundary operators involve shifts in both time and space depending on the order of the NRBC. Despite their accuracy improving as the order is increased, there comes a point where the intensive computation involved outweighs any benefit gained by proceeding to higher order.

Finally, one may analyse DNRBCs by studying the semi-discrete problem, obtained by discretising the spatial derivatives only. This approach was taken by Halpern [37] to analyse several DNRBCs for the wave equation. Her idea was to use Padé approximants to find improved boundary conditions that best matched the continuous NRBC for the wave equation, i.e. (2.1).

The preceding list of approaches to non-reflecting boundary conditions is best summarised by figure 2.1.

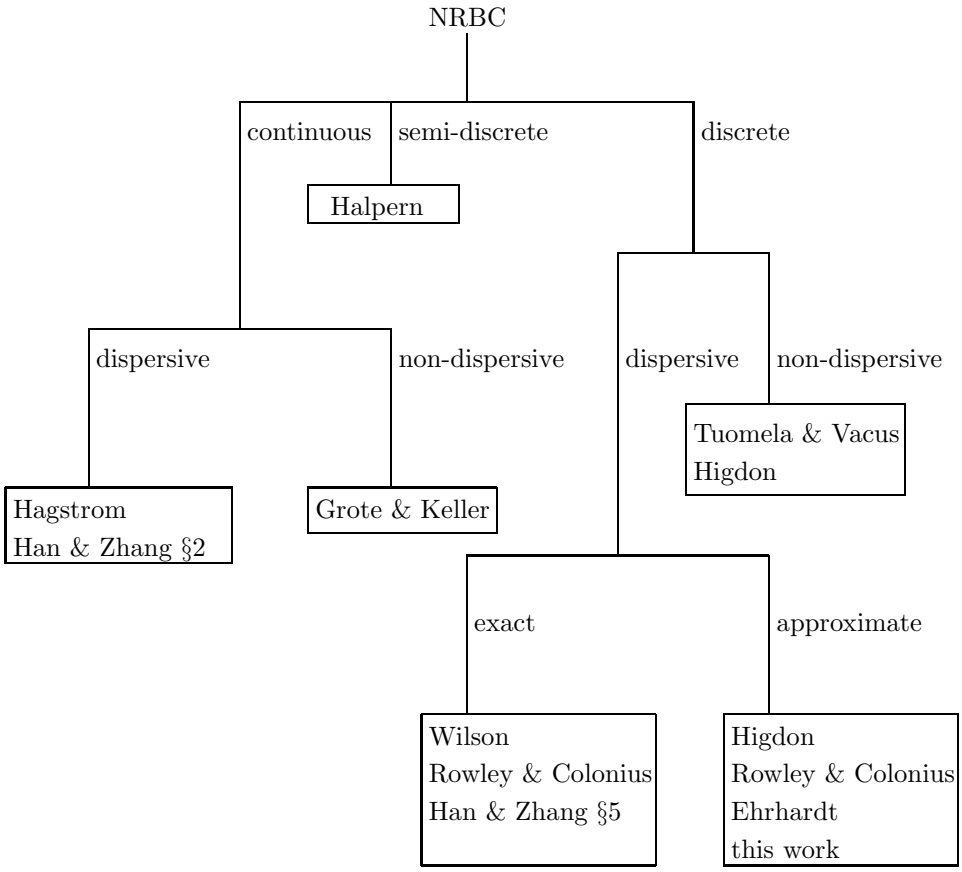


Figure 2.1: The hierarchy of NRBCs, as considered by a variety of authors.

2.3 Non-reflecting boundary conditions for a 1-d dispersive system

In this section, we consider the Klein-Gordon equation, being a relatively simple case of a dispersive equation, and choose like Higdon (for reasons that will become clear) to work with a second order space-/time-centred finite-difference scheme. We study one *exact* discretely non-reflecting (i.e. reflectionless) boundary condition for the Klein-Gordon equation in one dimension when solved using a space-centred leapfrog finite-difference scheme, as derived by Han and Zhang [38]. They compared this method with an exactly non-reflecting boundary condition obtained for the original PDE, which is then discretised (and implemented using a fast algorithm). As a reference solution, they used an exact analytical solution of the Klein-Gordon equation.

The choice of equation is a practical one: several problems in later chapters involve the coupling of a system to a "string" governed by the Klein-Gordon equation. The Klein-Gordon equation also provides an example of a system that is simple (single variable in one dimension) whilst also exhibiting dispersion. It also has resonance here because gravity waves in the linearised shallow water equations in 2-d may be shown to be exactly governed by a 1-d Klein-Gordon equation. Many of the ideas of this chapter may be equally well applied to equations that are more complex, or those that accurately describe a real-world model; in this sense, the equation is a useful prototype.

We explore some of the boundary condition's features and limitations when employed practically, and propose some possible improvements and remedies. The boundary condition under consideration, although local in space, involves a convolution using the full history of the solution at a single grid point. We compare a finite-history approximation to this scheme — where the series of convolution coefficients is simply truncated — to a scheme of the same form, where a finite number of coefficients have instead been chosen to be optimal, and derive a simple result governing the stability of such a scheme from GKS theory. We also investigate the effect of combining such boundary conditions with the Asselin filter commonly used to damp the computational mode inherent in the leapfrog scheme, and derive a novel boundary condition to take account of the filtering.

2.3.1 Motivation

In the following sections we present an evaluation of some of the practical limitations of an exactly non-reflecting boundary condition for the 1-d Klein-Gordon equation, as derived in [38], when it is actually applied. This exact DNRBC is non-local in time, involving a convolution that requires the full history of the solution to be known at a single grid point. Since this may be impractical for computations over long times, it is natural to truncate the convolution, since the coefficients decay in magnitude. However, we show that taking such a reasonable step as truncating leads to a boundary condition that is sub-optimal, and even unstable in some cases. We provide an alternative method of choosing the convolution coefficients, so that reflections are minimised for a given number of coefficients, and stability is guaranteed. Although the DNRBC in question applies only to the Klein-Gordon equation, this equation is a natural model equation to choose because it is a simple example of a dispersive system, and is well represented in the literature on NRBCs.

Even when the exactly non-reflecting boundary condition found in [38] is used without truncation, numerical experiments show that a small level of boundary reflection remains. This is because in practice the leapfrog scheme is used together with the three-level Asselin filter, in order to damp the spurious mode with wavenumber $\pi/\Delta x$. As a result, the finite-difference scheme is altered from the one from which the boundary condition was derived. The performance is worsened as the level of filtering is increased. Given the widespread use of the Asselin filter (e.g. in numerical weather prediction), its effect on the performance of accurate boundary conditions is worth consideration. Later we show that the problem can be overcome by deriving a new boundary condition that accounts for the time filtering.

2.3.2 An exactly non-reflecting boundary condition and optimal coefficients

Consider the one-dimensional Klein-Gordon equation,

$$u_{tt} - c^2 u_{xx} + \alpha u = 0, \tag{2.2}$$

where c is the wave speed and α controls the level of dispersion[1]. Note that it is possible without loss of generality to perform a scaling such that $\alpha = c = 1$. However for the sake of convenience when we come to explore other parameter regimes, we do not make this simplification.

A non-reflecting boundary condition that performs perfectly in the limit of increasingly fine resolution in time and space differencing may be derived in the following manner, as shown by Hagstrom [35] and Han & Zhang [38]. In order to relate the solution at the right-hand ($x = L$) boundary to that in the interior we require a relation between u and u_x at the end point — a so called Dirichlet-to-Neumann map. Taking the Laplace transform in time of (2.2) gives

$$\mathcal{L}\{u_{tt} - c^2 u_{xx} + \alpha u\} = \sigma^2 \tilde{u}(x,\sigma) - \sigma u(x,0) - u_t(x,0) - c^2 \sigma \tilde{u}_{xx}(x,\sigma) + \alpha \tilde{u}(x,\sigma) = 0, \tag{2.3}$$

[1] The u term in the Klein-Gordon equation is also commonly seen multiplied by a *squared* constant, but we do not follow this convention.

where tildes denote transformed quantities. Assuming the solution at the boundary is initially stationary and undisplaced, rearrangement of (2.3) yields

$$\tilde{u}_{xx} = \frac{\sigma^2 + \alpha}{c^2}\tilde{u}, \tag{2.4}$$

and upon integrating

$$\tilde{u}(x,\sigma) = \exp\left[-\frac{\sqrt{\sigma^2+\alpha}}{c}x\right] + \xi(\sigma). \tag{2.5}$$

Only the decaying solution is retained, which has the correct behaviour as $x \to \infty$. To derive a Dirichlet-to-Neumann map, (2.5) need only be differentiated and the Laplace transform inverted to give

$$u_x(L,t) = -\frac{1}{c}u_t(L,t) - \frac{\sqrt{\alpha}}{c}\frac{J_1(\sqrt{\alpha}t)}{t} * u(L,t). \tag{2.6}$$

Here $J_n(\cdot)$ is a Bessel function of the first kind of order n, and a star denotes a convolution, i.e.

$$f(x) * g(x) = h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)dx. \tag{2.7}$$

Consider discretising (2.2) by letting u_m^n be an approximation to $u(m\Delta x, n\Delta t)$, with a total of M grid points in the interior. The right-hand, non-reflecting boundary is located at $x = (M+1)\Delta x$. Discretising (2.6) using one-sided forward differences in time and space, and rectangular quadrature, we obtain an expression for the solution at the $(M+1)$th point

$$u_{M+1}^n = \frac{1}{c\Delta t + \Delta x}\left[c\Delta t\, u_M^n + \Delta x\, u_{M+1}^{n-1} - \sqrt{\alpha}\Delta t\Delta x \sum_{k=1}^{n}\frac{J_1(k\sqrt{\alpha}\Delta t)}{k}u_{M+1}^{n-k}\right]. \tag{2.8}$$

Note that the boundary condition (2.6) reduces to that for the wave equation for $\alpha = 0$ as expected. Otherwise, the convolution term requires that the entire history of the solution be known. In [38] the so-called FAST algorithm is proposed to reduce the computational cost of applying the exact NRBC. In practice one may take other steps to reduce such overheads; for example by working with a solution that is quiescent for $t < t_0$. One may also take the step of truncating the convolution, since, from the asymptotic form of the Bessel function for large argument, the coefficients decay like $k^{-3/2}$. When applied to a finite difference scheme, this NRBC is exact as $\Delta x, \Delta t \to 0$. However, to obtain a reflectionless scheme for finite resolution, we must turn to the finite-difference scheme first.

First, form a finite-difference approximation to (2.2) using the centred-time (leapfrog)/centred-space difference scheme

$$\frac{u_m^{n+1} - 2u_m^n + u_m^{n-1}}{\Delta t^2} - c^2\frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{\Delta x^2} + \alpha u_m^n = 0. \tag{2.9}$$

Again the non-reflecting boundary is the right-hand one, with the boundary point located at $m = M+1$. An exact DNRBC for this scheme can be derived using either the z-transform (which we will define), or the discrete Fourier transform (DFT). While entirely equivalent, certain parts of the analysis of the boundary condition are made more straightforward by a choice of one or the other; the DFT is better suited to analysis of the NRBC's dispersion properties, whilst the z-transform is more appropriate for examining its stability. The following

definition of the z-transform of a discrete series $x[n]$ is used in order to derive the discrete boundary condition:

$$f(z) = \mathcal{Z}\{x[n]\} = \sum_{n=0}^{\infty} x[n] z^{-n}. \tag{2.10}$$

Using this definition together with the shifting property of the z-transform, we transform (2.9) in time (i.e. in the index n) to obtain

$$\frac{1/z\,\tilde{u}_m - 2\tilde{u}_m + z\tilde{u}_m}{\Delta t^2} - c^2 \frac{\tilde{u}_{m+1} - 2\tilde{u}_m + \tilde{u}_{m-1}}{\Delta x^2} + \alpha \tilde{u}_m = 0, \tag{2.11}$$

where tildes denote transformed variables. Since (2.11) is a homogeneous, second order difference equation with constant coefficients, it has a solution of the form $\tilde{u}_m = A\tilde{\kappa}(z)^m$, where A is a constant. The function $\tilde{\kappa}(z)$ is obtained as the solution of

$$\frac{1/z - 2 + z}{\Delta t^2} - c^2 \frac{\tilde{\kappa} - 2 + 1/\tilde{\kappa}}{\Delta x^2} + \alpha = 0, \quad |\tilde{\kappa}(z)| < 1. \tag{2.12}$$

The condition on the magnitude of $\tilde{\kappa}$ ensures that $|\kappa_k| \to 0$ as $k \to \infty$, where κ_k is the kth coefficient in the series. Applying the general solution of the difference equation at the boundary point, we obtain $\tilde{u}_{M+1} = \tilde{\kappa}(z)\tilde{u}_M$ which when inverted using the convolution theorem for z-transforms gives the following NRBC:

$$u^n_{M+1} = \kappa_k * u_M = \sum_{k=0}^{n-1} \kappa_{n-k} u^k_M, \tag{2.13}$$

i.e. the boundary value at the nth time level, u^n_{M+1}, is given by a convolution involving the solution at the adjacent grid point at all previous times.

2.3.3 Inverting the z-transform

Consider the inversion procedure required to determine the coefficients κ_k. Solving (2.12) gives

$$\begin{aligned} \tilde{\kappa}(z) = \quad & 1 + \alpha\Delta x^2/2c^2 + \Delta x^2/c^2\Delta t^2(z - 2 + z^{-1}) \\ & \pm\sqrt{\left[1 + \alpha\Delta x^2/2c^2 + \Delta x^2/c^2\Delta t^2(z - 2 + z^{-1})\right]^2 - 1}. \end{aligned} \tag{2.14}$$

The choice of sign in front of the square root must be made to correspond to the outgoing solution; in this case this means taking the negative square root. One may take several approaches to inverting $\tilde{\kappa}(z)$, given that there are no standard results available for the square root term. One is to attempt to analytically evaluate the Bromwich integral which defines the inverse z-transform [46]

$$\kappa_k = \frac{1}{2\pi \mathrm{i}} \oint \tilde{\kappa}(z) z^{k-1} dz, \tag{2.15}$$

taking care to integrate along the branch cuts introduced by the square root. Another method is to use the relationship between κ_k and the Laurent series of $\tilde{\kappa}(z)$; a computer algebra system may then be used to compute this Laurent series exactly, making inversion relatively straightforward. In another approach, the function $\tilde{\kappa}(z)$ may actually be inverted exactly and explicitly, following the steps used in [20] to determine the exact DNRBC for the Schrödinger equation. That case is simpler since the equation is first order in time, but the process is the same.

Take the square root term in (2.14); one may write it as

$$\frac{1}{2z}\frac{\Delta x}{c^2\Delta t^2}\sqrt{(Az^2-2Bz+A)(Dz^2-2Ez+D)}, \tag{2.16}$$

where

$$A=1, \quad B=1-\alpha\Delta t^2/2, \quad D=\Delta x^2, \quad E=2\Delta x^2-2c^2\Delta t^2-\alpha\Delta x^2\Delta t^2/2. \tag{2.17}$$

Then rewrite the square root in the following way

$$\begin{aligned}\sqrt{Az^2-2Bz+A}\sqrt{Dz^2-2Ez+D} &= \frac{1}{\sqrt{A}}\frac{Az^2-2Bz+A}{z}\frac{z}{\sqrt{z^2-\frac{2B}{A}+1}} \\ &\quad\times\frac{1}{\sqrt{D}}\frac{Dz^2-2Ez+D}{z}\frac{z}{\sqrt{z^2-\frac{2E}{D}+1}} \qquad (2.18)\\ &= \frac{1}{\sqrt{AD}}\frac{(Az^2-2Bz+A)(Dz^2-Ez+D)}{z^2} \\ &\quad\times G(z,\mu_1)G(z,\mu_2), \qquad (2.19)\end{aligned}$$

where

$$G(z,\mu)=\frac{z}{\sqrt{z^2-2\mu\ +1}} \tag{2.20}$$

and

$$\mu_1=B/A, \quad \mu_2=E/D. \tag{2.21}$$

Hence the full square root term in (2.14) is given by

$$\begin{aligned}&\frac{\Delta x}{2c^2\Delta t^2\sqrt{AD}}\left[ADz-(2BD+2AE)+(2AD+4BE)z^{-1}\right. \\ &\qquad\left.-(2AE+2BD)z^{-2}+CDz^{-3}\right]G(z,\mu_1)G(z,\mu_2).\end{aligned} \tag{2.22}$$

Now, all parts of this expression other that the function G are trivially invertible. Fortunately, we may go further by noticing that $G(z,\mu)$ is the generating function of the Legendre polynomials in μ. Hence,

$$\mathcal{Z}^{-1}\left\{G(z,\mu)\right\}=\mathcal{Z}^{-1}\left\{\frac{z}{z^2-2\mu z+1}\right\}=P_n(\mu). \tag{2.23}$$

Using this result and the convolution theorem for z-transforms, together with (2.18) and (2.14), we obtain

$$\begin{aligned}\kappa_k = \quad &\Delta x^2/2c^2\Delta t^2\delta(k+1)+(1+\alpha\Delta x^2/2c^2-\Delta x^2/c^2\Delta t^2)\delta(k)+\Delta x^2/2c^2\Delta t^2\delta(k-1) \\ &+\frac{\Delta x}{2c^2\Delta t^2\sqrt{AD}}\left[AD\delta(k+1)-(2BD+2AE)\delta(k)+(2AD+4BE)\delta(k-1)\right. \\ &\left.-(2AE+2AD)\delta(k-2)+CF\delta(k-3)\right]*P_n(\mu_1)*P_n(\mu_2),\end{aligned} \tag{2.24}$$

where $\delta(k)$ is the discrete Dirac-delta function. This form of the coefficients is perhaps not the most helpful, involving as it does two convolutions, together with the Legendre polynomials which are themselves defined recursively.

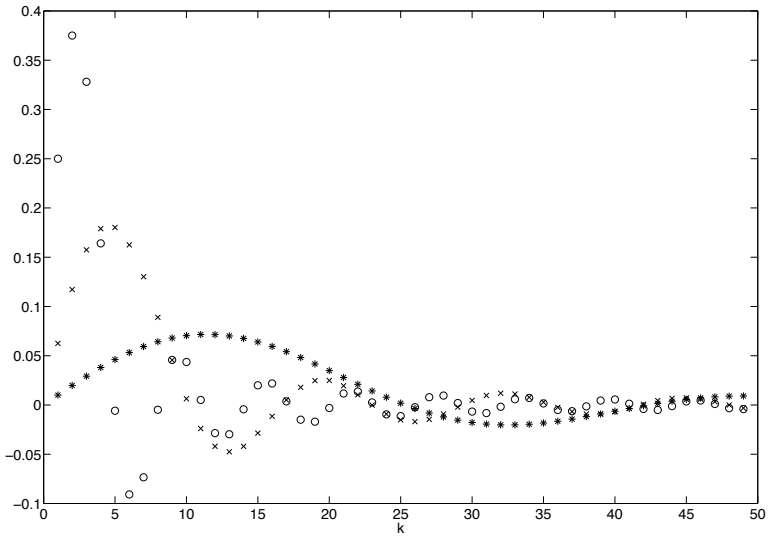


Figure 2.2: The first hundred DNRBC coefficients for three sets of parameter values: $\alpha = c = 1$ (circles); $\alpha = 2.5, c = 0.5$ (crosses); $\alpha = 5, c = 0.2$ (stars). In all three cases $\Delta x = 0.01$ and $\Delta t = 0.005$.

In terms of both convenience and accuracy, probably the best method for computing κ_k is to determine a recurrence relation by solving an ODE as shown in [38] (where the non-reflecting boundary was instead placed at the left-hand side of the domain, rather than the right-hand side as is done here). In our notation the coefficients are given by the following expressions:

$$\kappa_1 = \frac{\Delta x^2}{2c^2\Delta t^2}(1 - a_2), \quad \kappa_n = -\frac{\Delta x^2 a_{n+1}}{2c^2\Delta t^2},\ n \geq 2. \tag{2.25}$$

Here,

$$a_1 = \frac{r_1}{2}, \quad a_2 = \frac{r_2}{2} - \frac{a_1 r_1}{4}, \quad a_3 = \frac{r_1}{2} - \frac{a_2 r_1}{2}, \quad a_4 = \frac{1}{2} + \frac{a_1 r_1}{8} - \frac{a_2 r_2}{4} - \frac{5a_3 r_1}{8}, \tag{2.26}$$

and

$$a_n = \frac{1}{2n}\left[(3-2n)r_1 a_{n-1} + (6-2n)r_2 a_{n-2} + (9-2n)r_3 a_{n-3} + (12-2n)a_{n-4}\right], \tag{2.27}$$

where

$$r_1 = 2\alpha\Delta t^2 - 4 + \frac{4c^2\Delta t^2}{\Delta x^2}, \quad r_2 = -4\Delta t^2\alpha + 6 + \alpha^2\Delta t^4 + \frac{4\alpha c^2\Delta t^4}{\Delta x^2} - \frac{8\Delta t^2 c^2}{\Delta x^2}. \tag{2.28}$$

Some examples of the coefficients obtained are displayed in figure 2.2, for three different sets of parameter values. Figure 2.3 shows the algebraic decay of one such set of coefficients.

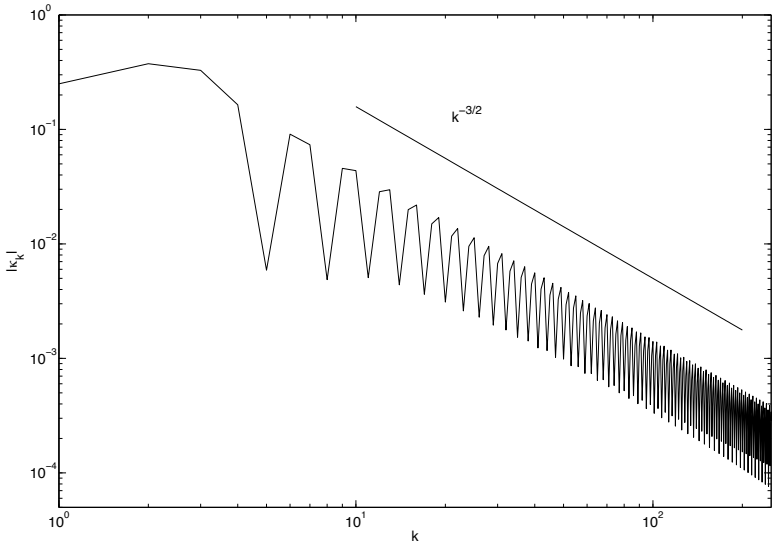


Figure 2.3: Decay in the magnitude of κ_k as $k \to \infty$, with $\alpha = c = 1, \Delta x = 0.01, \Delta t = 0.005$.

In this chapter, we employ a novel method for computing the coefficients κ_k, which is accurate and also simpler than other approaches. A convenient way of obtaining the coefficients $\kappa_1, \kappa_2, \ldots, \kappa_{n-2}, \kappa_{n-1}$, is to approximate them numerically using the discrete Fourier transform (DFT). The values obtained using this approach converge to the exact values of the coefficients as $N \to \infty$, where N is the number of sampling points. In practice we may choose another transform with similar shift properties, but the DFT (whose use for this purpose has not appeared in the literature) makes the process of computing the coefficients an automatic one. The DFT in time of u_m^n is given by

$$\hat{u}_m[l] = \sum_{n=1}^{N} \exp(-2\pi i l n/N) u_m^n, \tag{2.29}$$

and by substituting in (2.9) and using the shift property of the DFT we find

$$\frac{2\hat{u}_m(\cos(2\pi l/N) - 1)}{\Delta t^2} - c^2 \frac{\hat{u}_{m+1} - 2\hat{u}_m + \hat{u}_{m-1}}{\Delta x^2} + \alpha \hat{u}_m = 0. \tag{2.30}$$

This difference equation can be solved at the boundary point by letting $\hat{u}_{M+1} = \hat{\kappa}\hat{u}_M$, so that $\hat{\kappa}[l]$ is the solution of

$$-\frac{4}{\Delta t^2}\sin^2(\pi l/N) - c^2 \frac{\hat{\kappa} - 2 + 1/\hat{\kappa}}{\Delta x^2} + \alpha = 0. \tag{2.31}$$

Solving (2.31) for $\hat{\kappa}(l)$, choosing the positive root corresponding to right-going waves and performing the inverse DFT, we find that

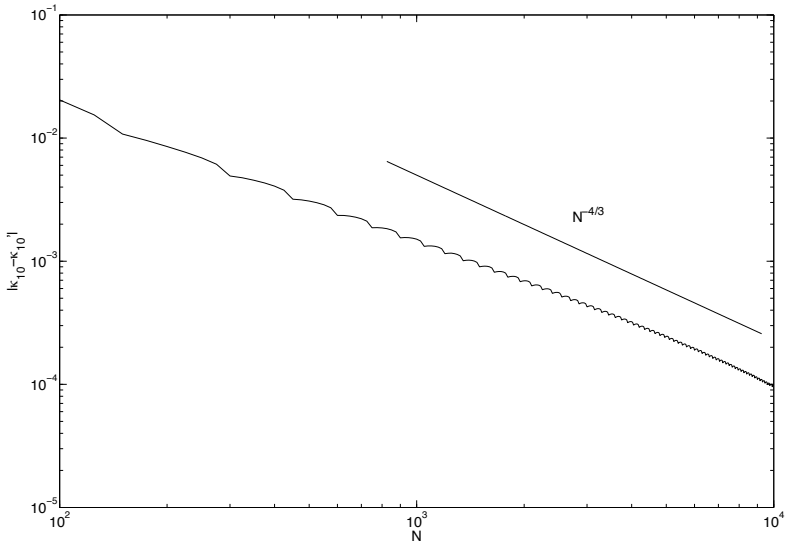


Figure 2.4: Error in the tenth DNRBC coefficient as computed by the DFT method, as the number of sampling points N is varied. A line whose slope corresponds to $N^{-4/3}$ is shown for comparison.

$$\kappa_k \approx \mathrm{Re}\left\{\frac{1}{N}\sum_{l=1}^{N}\exp(2\pi \mathrm{i} lk/N)\left[\frac{\Delta x^2}{c^2}\left(\frac{1}{\Delta t^2}\cos\left(\frac{2\pi\, l}{N}\right)-\frac{1}{\Delta t^2}+\frac{\alpha}{2}+\frac{c^2}{\Delta x^2}\right)\right.\right.$$
$$\left.\left.+\frac{1}{2}\sqrt{\left(\frac{\Delta x^4}{c^4}\right)\left(\frac{2}{\Delta t^2}\cos\left(\frac{2\pi\, l}{N}\right)-\frac{2}{\Delta t^2}+\alpha+\frac{2c^2}{\Delta x^2}\right)^2-4}\right]\right\}. \qquad (2.32)$$

Since the inverse DFT is periodic (i.e. symmetric about the midpoint) only the first $N/2$ values are valid coefficients. We require $N/\Delta x \gg 1$ in order to obtain a sufficiently accurate approximation. A noteworthy point overlooked in the literature is that convergence of such an approximation to the exact coefficients is not particularly rapid; we find that for a single coefficient, the error measure $|\kappa_k - \kappa'_k|$ — where the dash indicates a coefficient approximated by the DFT method — decays similarly to $N^{-4/3}$ as $N \to \infty$, as illustrated by figure 2.4. This is perhaps why methods employing either computer algebra to compute the Laurent series, or determination of a recurrence relation, are perhaps preferable.

The exactly non-reflecting boundary operator (2.13) requires the entire history of the solution at the grid point neighbouring the boundary. Such a procedure becomes computationally intensive as the number of time steps becomes very large. A fast strategy for evaluating the convolution may be required, and several such methods have been proposed by Antoine et al [2]. However, since the convolution will continue to become increasingly expensive for growing run times, it may be practical to truncate the convolution at some point, as investigated by Ehrhardt [20] for the Schrödinger equation. We consider such an approach when applied to the scheme of Han & Zhang. Of course, truncating the coefficient series will render the scheme approximate; the boundary condition is no longer exact and as such will cease to be reflectionless.

The NRBC coefficients decay relatively slowly($|\kappa_k|$ decays proportionally to $k^{-3/2}$ as shown in [20]) making it necessary to retain a fairly large number of terms for the convolution to remain accurate. We quantify the effect that reducing the number of coefficients to $K \leq n$ has on the reflection properties of the scheme as follows. Assume a solution to the difference scheme of the form $\exp(\mathrm{i}(n\omega\Delta t - m\xi\Delta x))$, where ω and ξ are the respective duals of t and x. Substituting in (2.9) we obtain the discrete dispersion relation of the finite-difference scheme in the implicit form

$$-\frac{2}{\Delta t^2}\sin^2(\omega\Delta t/2) + \frac{2c^2}{\Delta x^2}\sin^2(\xi\Delta x/2) + \frac{\alpha}{2} = 0; \quad -\pi \leq \xi,\ \omega \leq \pi. \tag{2.33}$$

We define the *reflection coefficient* R in the standard way (i.e. as in [23]) as the ratio of amplitudes of an incident and reflected wave at the boundary:

$$R = \frac{B(-\xi\Delta x, \omega\Delta t)}{B(\xi\Delta x, \omega\Delta t)}, \tag{2.34}$$

where B is the symbol of the boundary operator given in (2.13), about the point u_M^n:

$$B(\xi\Delta x, \omega\Delta t) = e^{\mathrm{i}\xi\Delta x} - \kappa_1 e^{-\mathrm{i}\omega\Delta t} - \kappa_2 e^{-2\mathrm{i}\omega\Delta t} - \cdots - \kappa_K e^{-K\mathrm{i}\omega\Delta t}. \tag{2.35}$$

By solving (2.33) for ω and substituting in (2.34), we obtain R as a function of ξ only. As an example, let us restrict the problem to choosing a single coefficient, κ_1 (a simplification we will often use for illustrative purposes). From (2.33) we find

$$\omega(\xi) = \frac{1}{\Delta t}\arccos\left(\frac{h(\xi)}{2\Delta t^2}\right), \tag{2.36}$$

where we have defined

$$h(\xi) = -2c^2\Delta t^2 + 2\Delta x^2 - \alpha\Delta t^2\Delta x^2 + 2c^2\Delta t^2\cos(\xi\Delta x). \tag{2.37}$$

Now substituting (2.35) into (2.34), making use of (2.36) and simplifying yields

$$r(\kappa_1) = \frac{\mathrm{e}^{\mathrm{i}\xi\Delta x} - \frac{h(\xi)}{2\Delta x^2} + \frac{\mathrm{i}}{2}\sqrt{4 - \frac{h(\xi)^2}{\Delta x^4}}}{\mathrm{e}^{-\mathrm{i}\xi\Delta x} - \frac{h(\xi)}{2\Delta x^2} + \frac{\mathrm{i}}{2}\sqrt{4 - \frac{h(\xi)^2}{\Delta x^4}}}. \tag{2.38}$$

Figure 2.5 shows the magnitude of the reflection coefficient as a function of ξ, for various truncation lengths with a particular choice of the simulation parameters. As K increases the curves become closer to the ideal zero over the range $\xi\Delta x \in [0, \pi]$. Note that there is always the constraint $R(0) = R(\pi) = 1$, but it turns out that this is not too constrictive, and is actually required. To understand the former point, consider the discrete group velocity for this discretisation of the Klein-Gordon equation:

$$C_{\mathrm{g}} = \frac{2\,c^2\Delta t\,\sin(\xi\,\Delta x)}{\Delta x\sqrt{4 - \frac{h(\xi)^2}{\Delta x^4}}}. \tag{2.39}$$

Hence $C_{\mathrm{g}}(0) = 0$, i.e. waves with zero wavenumber do not propagate inwards or outwards, and so the NRBC being totally reflecting for these modes is not deleterious. Also, for the boundary

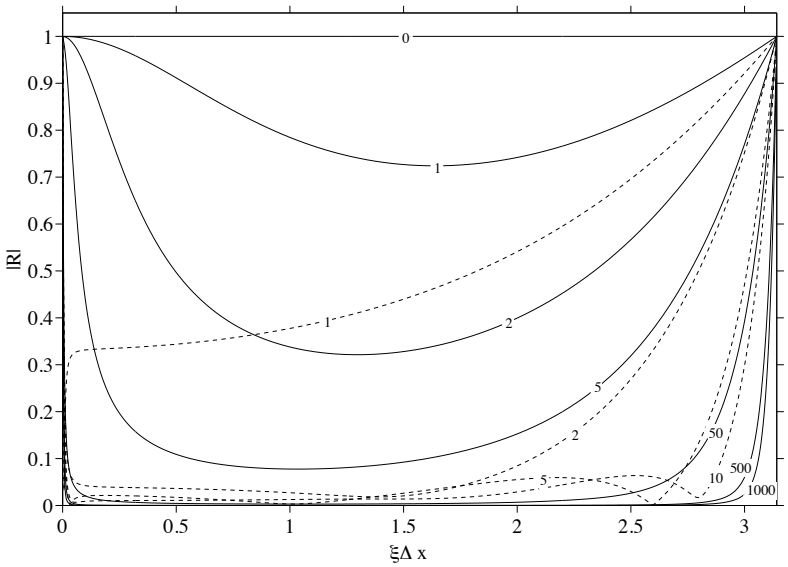


Figure 2.5: Magnitude of the reflection coefficient for various choices of convolution coefficients. The solid curves correspond to coefficient series obtained by truncating the exactly-non-reflecting series; in each case the number of coefficients retained is marked. The dashed lines correspond to optimal coefficient choices, where the marked number again indicates the number of coefficients. Other parameters were $\alpha = c = 1$, $\Delta x = 0.01$ and $\Delta t = 0.005$.

condition to be a non-reflecting one, it should be (at least in some approximate way) satisfied by outgoing waves only, and not satisfied by those that are incoming. This differing treatment of the two types of wave cannot extend to the case with $\xi = 0$, as we have just seen that these waves neither leave nor enter the domain. Hence, the boundary condition cannot distinguish between incoming and outgoing behaviour in this case, and so the reflection coefficient must be unity.

We can define a measure of the effectiveness of the boundary conditions with different levels of truncation, by defining

$$R_{\text{int}} = \frac{1}{\pi} \int_0^\pi |R(\xi \Delta x)| d\xi \Delta x. \tag{2.40}$$

A perfectly non-reflecting scheme corresponds to $R_{\text{int}} = 0$, whereas a hard wall condition (by virtue of the normalising factor of π) corresponds to $R_{\text{int}} = 1$. The choice of the measure (2.40) is an arbitrary one, and measures of different forms could be investigated. Figure 2.6 shows how R_{int} (which is calculated numerically) tends to zero approximately algebraically as the number of coefficients is increased. One observation is that by minimising the L_1 norm, we effectively give equal weight to reflections of all possible wavenumbers, whereas for example, the L_2 norm may favour the minimisation of reflections with extremal wavenumbers.

Simply truncating the coefficient series, which was originally derived to be perfectly reflectionless, is one method of forming a practical boundary condition. However this strategy does not guarantee that the coefficients will be the best possible choice. In addition to being sub-optimal, such a scheme can even be unstable when only a handful of coefficients are retained. As an example, for the parameter values already investigated, the boundary value blows up

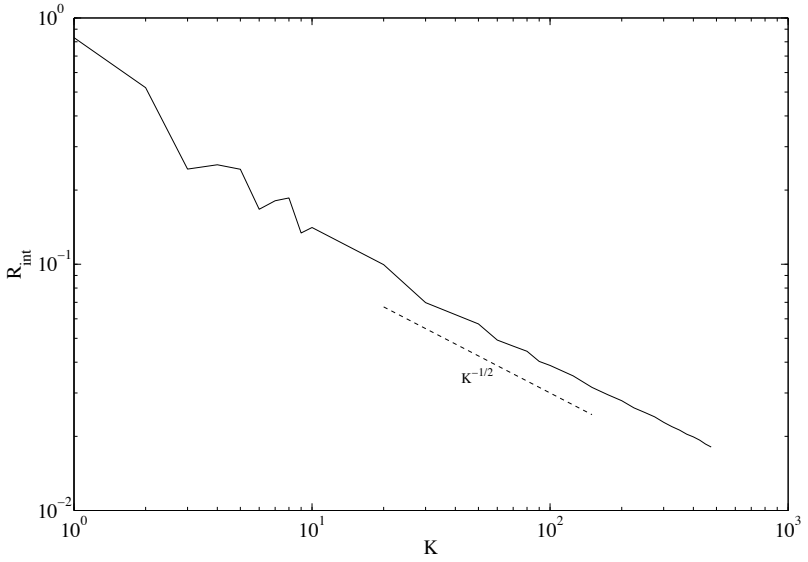


Figure 2.6: Log-log plot of R_{int} as a function of the number of convolution coefficients K.

exponentially when $K = 4, 5, 6$ and even for $K = 23$. Instead of truncating the coefficient series, an entirely new approach is taken: the coefficients are chosen to be the solution of a minimisation problem, with the objective function being R_{int} (which is evaluated numerically) for a particular choice of the problem parameters. The optimal coefficients for a particular choice of simulation parameters were found using the Nelder-Mead simplex search algorithm, implemented in the `fminsearch` function in MATLAB. The truncated coefficient series was taken as a starting point for the optimisation for each set of K coefficients, and values sought that would minimise R_{int}. Table 2.1 shows the coefficients obtained using this approach for one to fifteen coefficients, together with R_{int} obtained using these optimal choices and by truncation of the original perfectly absorbing coefficient series. Figure 2.5 also shows some examples of $|R|$ for the optimal coefficients. Note how careful choice of the coefficients can introduce nulls into the reflection coefficient, rendering the boundary perfectly non-reflecting for a particular wavenumber. Notice also, from the table, how for small K, R_{int} is dramatically reduced by using the set of optimal coefficients. It is interesting to note that despite the same truncated series being used as an initial starting point, there seems to be no obvious pattern behind the optimal values of κ_k as K is varied.

Figure 2.7 shows a comparative plot of the values of R_{int} obtained for both the truncated and the optimal sets of coefficients, for K up to fifty. Notice the similar rate of improvement in both schemes as the number of coefficients increases.

2.3.4 Stability for the boundary condition

Gustafsson, Kreiss and Sundström [32] developed a general stability theory (often referred to as the GKS theory) for mixed initial boundary values problems, and the difference approximations and non-reflecting boundary conditions used to solve them. A useful viewpoint on the GKS

K	κ_1	κ_2	κ_3	κ_4	κ_5	κ_6	κ_7	κ_8	κ_9	κ_{10}	κ_{11}	κ_{12}	κ_{13}	κ_{14}	κ_{15}	R_{int} (optimal)	R_{int} (truncated)
1	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5205	0.8318
2	−0.1616	1.1595	-	-	-	-	-	-	-	-	-	-	-	-	-	0.1644	0.5209
3	−0.1021	0.8834	0.2195	-	-	-	-	-	-	-	-	-	-	-	-	0.1517	0.2434
4	0.7080	−1.1401	2.2716	−0.8400	-	-	-	-	-	-	-	-	-	-	-	0.0817	0.2534*
5	0.8571	−1.6459	3.0118	−1.4246	0.2017	-	-	-	-	-	-	-	-	-	-	0.0770	0.2492*
6	0.2495	0.9654	−1.9520	3.8365	−2.9385	0.8406	-	-	-	-	-	-	-	-	-	0.0630	0.1671*
7	0.6564	−0.7741	1.3433	0.3441	−0.8464	0.2671	0.0102	-	-	-	-	-	-	-	-	0.0696	0.1810
8	0.5748	−0.4468	0.9550	0.1437	0.2055	−1.0027	0.7397	−0.1684	-	-	-	-	-	-	-	0.0645	0.1858
9	0.5198	−0.2756	0.7174	0.4234	−0.2925	−0.1661	−0.1530	0.3764	−0.1494	-	-	-	-	-	-	0.0613	0.1339
10	0.1726	0.7273	−0.1618	0.0151	0.9962	−0.6450	−0.9628	1.3340	−0.4687	−0.0078	-	-	-	-	-	0.0565	0.1409*
11	0.1650	0.9381	−0.8266	0.8703	0.6595	−1.0992	−0.2428	0.9161	−0.3655	−0.0269	0.0116	-	-	-	-	0.0533	0.1457*
12	0.1890	0.7092	−0.2053	0.3425	0.0786	0.5093	−1.1915	0.0952	1.2036	−0.8635	0.0143	0.1194	-	-	-	0.0475	0.1138
13	−0.4911	3.5457	−5.1330	4.2196	0.1156	−2.0690	0.5247	0.0186	1.0969	−1.0707	−0.0250	0.4147	−0.1481	-	-	0.0401	0.1212
14	−0.2588	2.4498	−3.0150	2.4814	−0.2826	−0.0144	−0.8900	−0.0471	1.2963	−0.5408	−0.5473	0.3147	0.1848	-0.1320	-	0.0386	0.1261
15	−0.3194	2.7316	−3.5326	2.8082	0.0800	−0.8801	−0.2339	−0.2784	1.5948	−1.1861	0.0894	0.0441	0.1792	−0.0884	−0.0095	0.0389	0.1010

Table 2.1: Optimal convolution coefficients derived by minimising R_{int}. Comparison is also made with the coefficients obtained by truncating the original coefficient series. Asterisks indicate cases in which the truncated boundary condition is unstable.

25

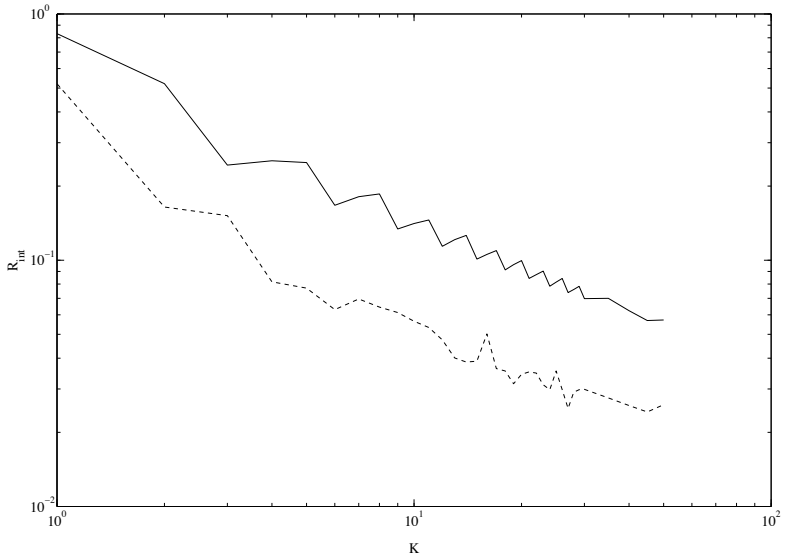


Figure 2.7: Log-log plot of R_{int} as a function of the number of convolution coefficients K, for the exactly non-reflecting coefficients after truncation (solid line), and the optimal coefficients (dotted line).

theory is provided by Trefethen [72], where the stability criterion is given a physical interpretation involving group velocity. Examples of application of the theory include [8], and papers by Higdon [39, 40], and §6.2 of [41]. The latter three references are useful as they include analyses of the finite-difference scheme which we have employed, together with Higdon's boundary conditions which generally involve several grid points in both time and space. In this section, we probe the stability of the system to varying degrees.

Numerical experiments show that if we follow the simple approach of truncating the exactly non-reflecting convolution coefficients, the resulting boundary condition can lead to instability, where the boundary value blows up to $-\infty$. Figure 2.8 illustrates the phenomenon. We seek a condition that guarantees the scheme is free of this type of problem. By applying separation of variables in the discrete setting, we consider solutions to (2.9) of the form

$$u_m^n = s^m z^n; \quad s, z \in \mathbb{C} \tag{2.41}$$

By substitution in (2.9) we obtain another form of the dispersion relation

$$z - 2 + z^{-1} = -\alpha \Delta t^2 + \left(\frac{c\Delta t}{\Delta x}\right)^2 (s - 2 + s^{-1}). \tag{2.42}$$

This equation has two roots s_1 and s_2 for each value of z. We also have the boundary operator D representing our boundary condition:

$$B(s, z) = s - \sum_{k=1}^{K} \kappa_k z^{-k}. \tag{2.43}$$

Stability of non-reflecting boundary conditions for hyperbolic initial-value problems is estab-

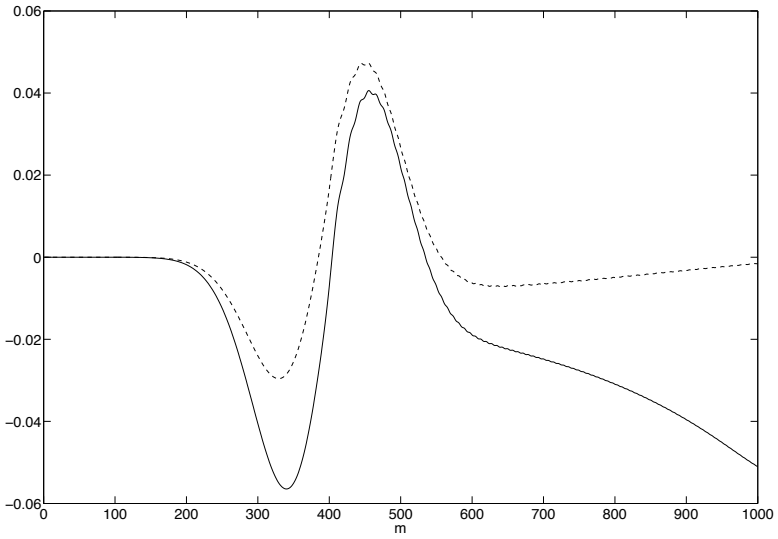


Figure 2.8: Stable and unstable solutions of the model equation, with two different sets of non-reflecting boundary conditions. The dashed line shows the stable solution, at the end point of which the value is beginning to settle down. On the other hand, the boundary value of the solution drawn with the solid line wanders — gradually at first — away from its resting value, never to return.

lished using the following criterion:

Theorem (Gustafsson Kreiss Sundström)**.** *A necessary and sufficient condition for the stability of an outflow boundary condition for a hyperbolic system is that*

$$B(s_l, z) \neq 0, \quad |z| > 1, \tag{2.44}$$

where s_l is the root of the discrete dispersion relation that corresponds to a left-going wave.

In other words, the condition asserts that there are no modes supported by the finite-difference scheme that travel into the domain from the boundary (i.e. left-going modes) that grow in time, and that also satisfy the boundary condition.

Consider the case of $z \in \mathbb{R}$ corresponding to non-oscillatory unstable modes as found in numerical experiments. This restriction of the problem simplifies the analysis somewhat, the general case of $z \in \mathbb{C}$ being considered later. It is necessary to classify the nature of the two roots of (2.42) for $z > 1$. To do so, we follow the line taken in §4 of [40], although the analysis is altered slightly by the presence of the $-\alpha\Delta t^2$ term in the discrete dispersion relation.

It can be shown that if $z > 1$, the two roots of (2.42) are real and furthermore $s_1 < 1$, and $s_2 > 1$. Consider oscillatory waves only:

$$\sigma = \mathrm{e}^{\mathrm{i}\xi\Delta x}, \quad z = \mathrm{e}^{\mathrm{i}\omega\Delta t}. \tag{2.45}$$

Substitution in the dispersion relation $D(\sigma, z)$ and subsequent rearrangement yields the relation

$$\sin^2(\omega\Delta t/2) - \alpha\Delta t^2/2 = (c\Delta t/\Delta x)^2 \sin^2(\xi\Delta x/2). \tag{2.46}$$

Before proving the statement about the locations of the roots of (2.42), we must first restate some of their general properties. First, the two roots s_1 and s_2 are the reciprocal of one another, since the dispersion relation is left unchanged by the transformation $z \to z^{-1}$. Now, consider $z > 1, z \in \mathbb{C}$: a glance at figure 2.9 is enough to establish that $\mathrm{Im}\{z - 2 + z^{-1}\} > 0$, and from (2.42), it is clear that neither of the roots s_1 and s_2 can be real or lie on the unit circle, since then the right hand side of the dispersion relation would be real (remembering that α is a real constant). Also as with both z_1 and z_2, $s_1 = s_2^{-1}$, so one root resides inside the unit circle, and the other outside.

Further, let us consider the case of z real, and prove the original statement. If $z > 1$, then the left hand side of the dispersion relation $z - 2 + z^{-1} + \alpha\Delta t^2/2$ is positive; this is clear since for $z > 0$ the only solution of $z - 2 + z^{-1} = 0$ is $z = 1$, and $\alpha > 0$. Hence, $s_1 - 2 + s_2^{-1} > 0$, and furthermore both roots must be real, as is made clear by figure 2.9. Since it was shown previously that only one root resides within the unit circle, it is clear that one of the roots may be found on the interval $(0, 1)$, and the other on $(1, \infty)$.

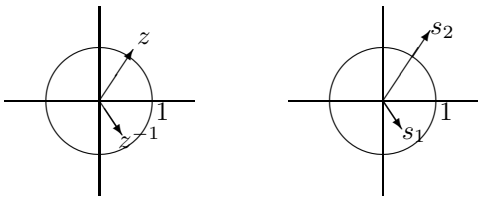


Figure 2.9: Positions of the roots of the discrete dispersion relation, for a general value of z.

Having classified the roots of the dispersion relation in general terms, it remains to derive the GKS condition. The root $s_2 > 1$ corresponds to a wave travelling into the domain for $z > 1$, as a glance at figure 2 in [72] confirms. In order to find a stability condition, it is required to fix z and solve (2.42) for s_2. Substituting these values into (2.44) gives

$$s_2 \neq \frac{\kappa_1}{z} + \frac{\kappa_2}{z^2} + \cdots + \frac{\kappa_{K-1}}{z^{K-1}} + \frac{\kappa_K}{z^K}, \tag{2.47}$$

where

$$s_2 = z - 2 + z^{-1} + \alpha\Delta t^2 + 2c^2\Delta t^2/\Delta x^2 + \frac{\sqrt{[-(z - 2 + z^{-1} + \alpha\Delta t^2) - 2c^2\Delta t^2/\Delta x^2]^2 - 4(c\Delta t/\Delta x)^4}}{2(c\Delta t/\Delta x)^2}. \tag{2.48}$$

Note that s_2 corresponds to the positive square root. Equation (2.47) holds for all $z \geq 1$ provided

$$s_2|_{z=1} > \sum_{k=1}^{K} \kappa_k, \tag{2.49}$$

Hence, we have the following new result.

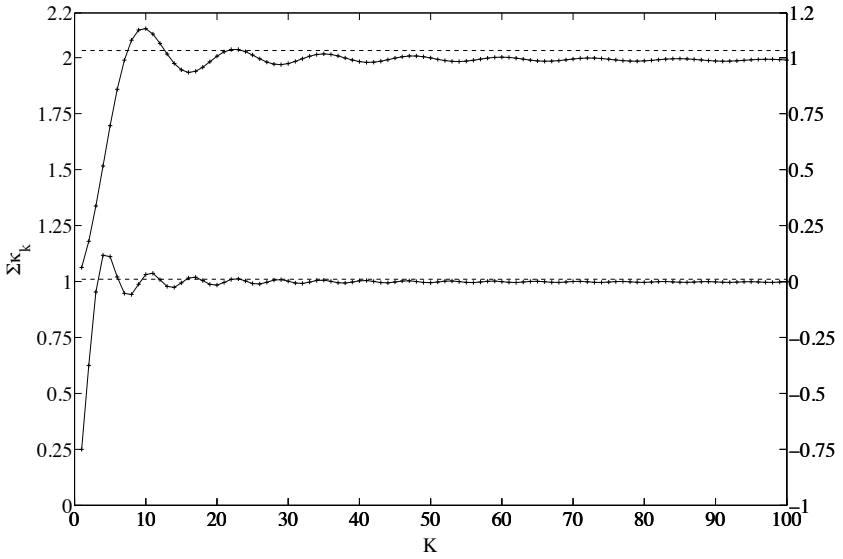


Figure 2.10: Plots showing the sum of the convolution coefficients (crosses on solid line) as the truncation length is varied, and the critical stability value (dashed line). Bottom curve (corresponding to left y-axis): $\alpha = c = 1$; top curve (corresponding to right y-axis): $\alpha = 2.5$, $c = 0.5$. In both cases $\Delta x = 0.01$ and $\Delta t = 0.005$.

Theorem. *A solution $u_m^n = u(m\Delta x, n\Delta t)$ to the 1-d linear Klein-Gordon equation, advanced using standard second-order centred finite differences and a discretely non-reflecting boundary condition of the form*

$$u_{M+1}^n = \kappa_k * u_M = \sum_{k=0}^{n-1} \kappa_{n-k} u_M^k, \tag{2.50}$$

is ***GKS-stable with respect to non-oscillatory growing modes****, provided that*

$$s_2|_{z=1} > \sum_{k=1}^{K} \kappa_k, \tag{2.51}$$

where s_2 is the largest root of the dispersion relation.

This constraint on the coefficients κ_k — a sufficient condition for stability — is particularly revealing when considering why the boundary condition obtained by truncation of the exact series is sometimes unstable. For example, experiments using values $\alpha = c = 1$, $\Delta x = 0.01$ and $\Delta t = 0.005$, show that the scheme is stable for all values of K except $K = 4, 5, 6, 10, 11, 16, 17$ and 23. A calculation shows that for this parameter choice $s_2|_{z=1} = 1.0101$. Figure 2.10(a) shows a plot of $\sum_{k=1}^{K} \kappa_k$ as the truncation length is varied: the value of the sum crosses the boundary between stable and unstable regions eight times before finally settling down in the stable region, thus explaining why the scheme is stable for all but a handful of values of K.

It is interesting to examine how stability is affected by the problem parameters. An illustration of the stability behaviour for another parameter choice is given in figure 2.10(b). Figure 2.3.4 shows how the critical stability value s_2 varies with both c and α. It is easy to show that the critical value for stability, $s_2|_{z=1}$, increases as $c \to 0$ and $\alpha \to \infty$, and so a choice of

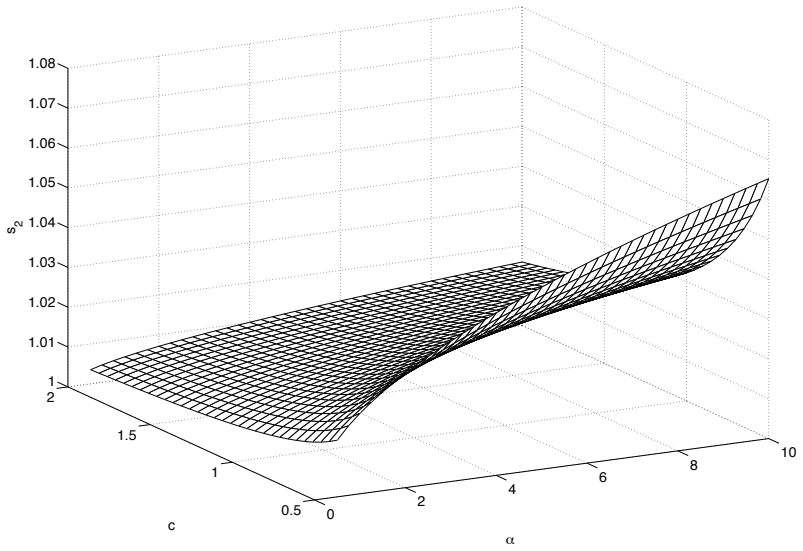


Figure 2.11: Surface showing the relation between the stability threshold $s_2|_{z=1}$, the wavespeed c and dispersion parameter α.

$c = 0.5$ and $\alpha = 2.5$ has been made to give $s_2|_{z=1} = 1.03213$. The figure shows that the value of K required to preclude the existence of these non-oscillatory instabilities, varies with the problem parameters. Taking this argument to its conclusion, it is possible in some situations that even with a large choice of truncation length K, which one might naively assume to give good results, the boundary condition may nevertheless be unstable.

We have not given conditions on κ_k that preclude the existence of unstable oscillatory modes; however, such instability was not observed with any of the coefficient series derived in the present work. It is however possible to construct coefficient sequences that yield other modes of instability. For example, choosing a particularly wild set of coefficients, it is possible to discover a sawtooth instability, corresponding to a wave of wavelength $2\Delta x$. Figure 2.3.4 shows an example of such behaviour.

The condition given above is sufficient but not necessary for stability of the scheme, allowing as it does for the possibility of the existence of exponentially growing oscillatory modes. In an effort to be more general and guarantee stability of the scheme using a particular set of coefficients, a simple numerical check may be carried out by using a Nyquist stability argument [63]. By solving (2.42) for s_l in terms of z and substituting in (2.35) we obtain an expression for the boundary operator, $g(z)$. Stability may be guaranteed by the GKS theorem, corresponding to demanding that $g(z)$ has no zeros outside the unit circle, and thus we use the principle of the argument to determine their placement. We examine the winding number I of the image in the complex plane of $g(z(t))$, where $z(t) = R\exp(2\pi \mathrm{i} t)$, and choose contours with $R = 1$ and $R > 1$. The total order of all the poles enclosed by both contours will be the same since the poles of $g(z)$ lie at the origin, so any difference in winding number will give us information as to the location of its zeros.

As an example, we choose the simple case of a single coefficient, and compare $\kappa_1 = 1.01$

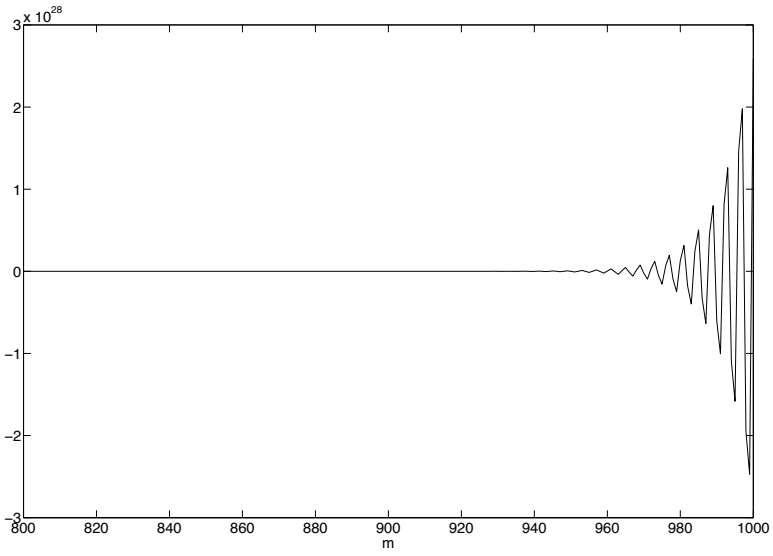


Figure 2.12: An unstable sawtooth mode makes its presence known. The coefficients do not satisfy (2.51).

and $\kappa_1 = 1.011$, which are known from the previous result to lie just either side of the stability threshold. We also take $\alpha = c = 1$, $\Delta x = 0.01$ and $\Delta t = 0.005$. Figure 2.13(a) shows the curve described in the complex plane by the image of $g(z)$ around the unit circle for $\kappa_1 = 1.01$, and at this scale the curve for $\kappa_1 = 1.011$ is much the same. For our larger contour we pick $R = 2$, which gives a curve (not shown) with $I = 0$. Figures 2.13(b) and (c) show $g(z(t))$ for $\kappa_1 = 1.01$ and 1.011, respectively. Together with figure 2.13(a) we see that for $\kappa_1 = 1.01$, $I = 0$, while for $\kappa_1 = 1.011$, $I = -1$, leading to the conclusion that in the second case a zero now lies outside the unit circle, rendering the scheme unstable.

Having derived a stability condition for the NRBC coefficients (2.44), it is now possible to ensure that coefficients generated by the optimisation procedure described in §2.3.2 will always yield a stable boundary condition. It is merely a question of adding (2.44) *as a constraint when performing the optimisation.* This is an approach which, from the literature, has yet to be tested.

2.3.5 Numerical results

We begin by examining how the performance of the exact DNRBC is affected by truncation of the convolution coefficients. Before doing so we must define a measure of the error. Han and Zhang [38] compared solutions obtained with their NRBCs with analytical solutions to the Klein-Gordon equation. Doing so does not highlight the effectiveness of the exact DNRBC, as discretisation error will always remain. Instead, rather than employing the solution to the original PDE, we use a reference solution computed on a domain large enough that no waves are reflected from the (fixed) boundary during the time of the experiment. The effectiveness of the boundary condition is then measured by comparing the error performance of the scheme with the DNRBC with that obtained on a domain of the same size, but using a hard wall condition

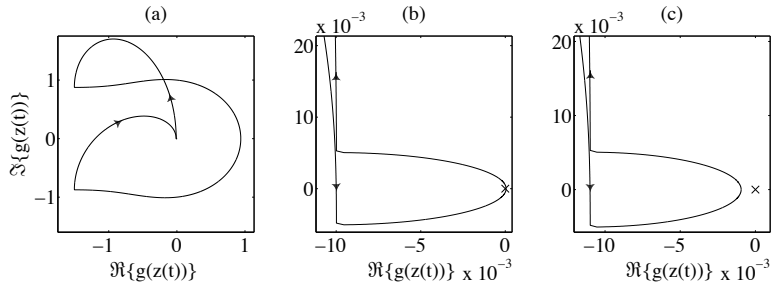


Figure 2.13: Images in the complex z-plane of the boundary operator for a single coefficient. Panels (a,b): $\kappa_1 = 1.01$; panel (c): $\kappa_1 = 1.011$. The origin is marked with an x in the "close-up" plots, and arrows mark the direction as t is varied from 0 to 2π. In (b) the curve only just misses enclosing the origin, whereas in (c) the curve winds once around zero.

($u^n_{M+1} = 0 \quad \forall n$). Specifically:

$$r(t) = \sqrt{\frac{\int_0^L (u - u_{\text{ref}})^2 dx}{\int_0^L (u_{\text{hard}} - u_{\text{ref}})^2 dx}}. \tag{2.52}$$

The maximum error at the boundary was also measured using the diagnostic

$$\max_t |u_{M+1} - u_{\text{ref},M+1}|. \tag{2.53}$$

All of the experiments were carried out using the parameter values $\alpha = c = 1$, $\Delta x = 0.01$, $\Delta t = 0.005$, with 200 grid points for the standard solution and 3000 for the reference solution, over 1000 time steps. The initial conditions used were

$$u(x, t = 0) = -\frac{1}{2\pi}\, x\, \exp(-5x^2), \quad u_t(x,\, t = 0) = 0. \tag{2.54}$$

An Asselin filter would commonly be used with the leapfrog scheme to damp the computational mode, however for reasons explained in §2.3.6, this was not done in these experiments.

First of all we consider truncating the original coefficient series. Figure 2.14 shows the behaviour of the reflection measure r at a particular time step, as the truncation length K is varied. In this scenario it takes just under 200 time steps for the wave to reach the boundary. We observe that for early times, as long as the difference in time steps between the wave striking the boundary and the current time step is less than K, the scheme is essentially exactly non-reflecting, with remaining small errors due to round-off. Beyond this, if we look at a fixed time (a single curve) increasing K leads to a decrease in the reflection error. The errors also grow

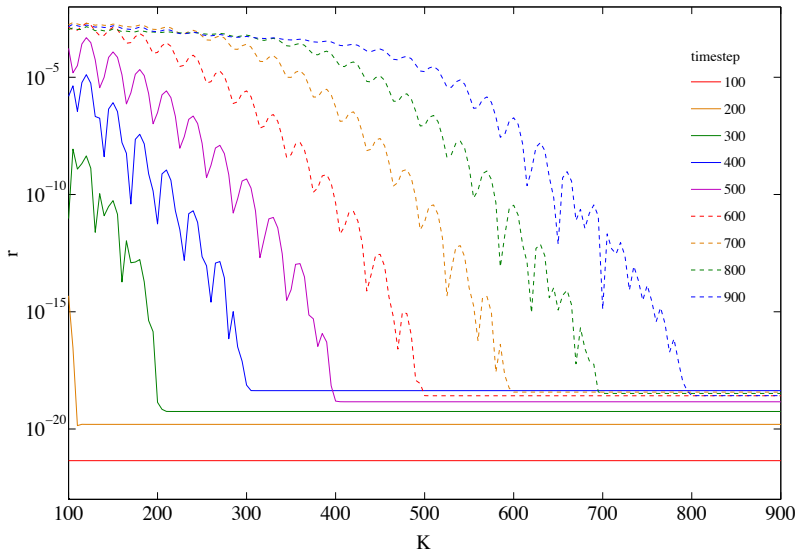


Figure 2.14: Performance of the truncated, exact DNRBC: each curve corresponds to a particular time step, while truncation length runs along the x-axis.

in time, when looking at fixed K.

Next we test the performance of the optimal coefficients, using the same experimental setup. Table 2.2 lists the maximum (over time) squared end point errors for the scheme, comparing the performance of the optimal coefficients and that of the truncated coefficient series. As an example of the improvement gained using the new method, for $K = 4$ we obtain errors using the optimal coefficients which may be acceptable in some scenarios, as opposed to a scheme which is not even stable using the truncated series. Although the optimal coefficients are seen to outperform the truncated series, one disappointing feature is the lack of any improvement for $K > 4$; it is unclear why this should be the case, given that R_{int} shows a steady improvement as the number of optimal coefficients is increased. Of course, many variations could be applied to improve this method and possibly alleviate this problem: more grid points at both the current and earlier times could be used to form the boundary operator, and one could choose another measure of the reflection coefficient as an objective function in the minimisation procedure. It must be borne in mind that the improvement comes at the computational cost of determining the optimal coefficients, though because they are computed once and for all for a given choice of parameters this is often a price worth paying.

2.3.6 Asselin filter effects

The boundary conditions studied here have been derived from consideration of the specific finite-difference scheme used to solve the problem. As soon as that scheme is altered, the performance of these boundary conditions may be adversely affected. Such degradation of results would be particularly noticeable in the case of the exact DNRBC without truncation, which is essentially perfectly absorbing (their error is zero to the level of machine accuracy).

	optimal		truncated	
K	max. end point error	$r(500\Delta t)$	max. end point error	$r(500\Delta t)$
1	1.395227×10^{-2}	3.069543×10^{-1}	4.726204×10^{-2}	9.968286×10^{-1}
2	3.837619×10^{-3}	7.558349×10^{-2}	4.674608×10^{-2}	9.743700×10^{-1}
3	5.741999×10^{-3}	1.138212×10^{-1}	2.412219×10^{-2}	4.523817×10^{-1}
4	1.596923×10^{-3}	2.583054×10^{-2}	-	-
5	1.631854×10^{-3}	3.400842×10^{-2}	-	-
6	2.824623×10^{-3}	6.140525×10^{-2}	-	-
7	1.749907×10^{-3}	4.309125×10^{-2}	2.814156×10^{-2}	5.042693×10^{-1}
8	1.921756×10^{-3}	4.601209×10^{-2}	2.979069×10^{-2}	5.389775×10^{-1}
9	1.670614×10^{-3}	4.054248×10^{-2}	7.971121×10^{-3}	1.354237×10^{-1}
10	1.772065×10^{-3}	3.056175×10^{-2}	-	-
11	1.639737×10^{-3}	3.268946×10^{-2}	-	-
12	2.556520×10^{-3}	5.575743×10^{-2}	8.081933×10^{-3}	1.637612×10^{-1}
13	2.231587×10^{-3}	4.216294×10^{-2}	1.399240×10^{-2}	2.285433×10^{-1}
14	2.033092×10^{-3}	3.715026×10^{-2}	1.660996×10^{-2}	2.719847×10^{-1}
15	2.081802×10^{-3}	3.772727×10^{-2}	4.512723×10^{-3}	6.632678×10^{-2}

Table 2.2: Performance comparison figures for the non-reflecting scheme using optimal and truncated coefficients. The dashes mark values of K for which truncating the exact coefficient series yields an unstable boundary condition. Again the parameter choices are $\alpha = c = 1$, $\Delta x = 0.01$, $\Delta t = 0.005$. The simulation was run over 1000 time steps on a 200 point domain.

Asselin [3] proposed a filter — subsequently studied by Robert [66] — that can be used with the leapfrog finite-difference scheme to damp the computational mode. However, typically NRBCs are derived by neglecting such a filtering procedure — in this work, we take account of it.The filtering procedure involves an update to the solution at the nth time-step of the form:

$$\bar{u}^n = u^n + \gamma(u^{n+1} - 2u^n + \bar{u}^{n-1}), \tag{2.55}$$

where bars denote filtered values. With filtering our finite-difference scheme takes the following revised form:

$$\frac{u_m^{n+1} - 2u_m^n + \bar{u}_m^{n-1}}{\Delta t^2} - c^2 \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{\Delta x^2} + \alpha u_m^n = 0. \tag{2.56}$$

Typical values for γ lie in the range 0.01–0.3. By adding time filtering to our scheme our boundary condition derived for pure leapfrog is no longer valid. To test the effect of the filter on the performance of the DNRBC we again run the experiment, this time over 1000 time steps without truncating the boundary coefficients. Under such conditions and without filtering, the boundary condition would be expected to be perfectly non-reflecting, with any remaining errors due to numerical round-off. The parameter γ is varied in the range 0–0.3 in the experiment. Figure 2.15 shows that the quality of the computed solution deteriorates markedly as γ is increased. We observe the same problem with the coefficients computed to be optimal, also shown in figure 2.15.

In an effort to counter such effects, we derive a new DNRBC from (2.55) and (2.56), which accounts for the filtering process. Taking the DFT in time of (2.55) and using the shift property yields

$$\hat{\bar{u}}_m = \hat{u}_m + \gamma(e^{2\pi \mathrm{i} l/N}\hat{u}_m - 2\hat{u}_m + \mathrm{e}^{-2\pi \mathrm{i} l/N}\hat{\bar{u}}_m), \tag{2.57}$$

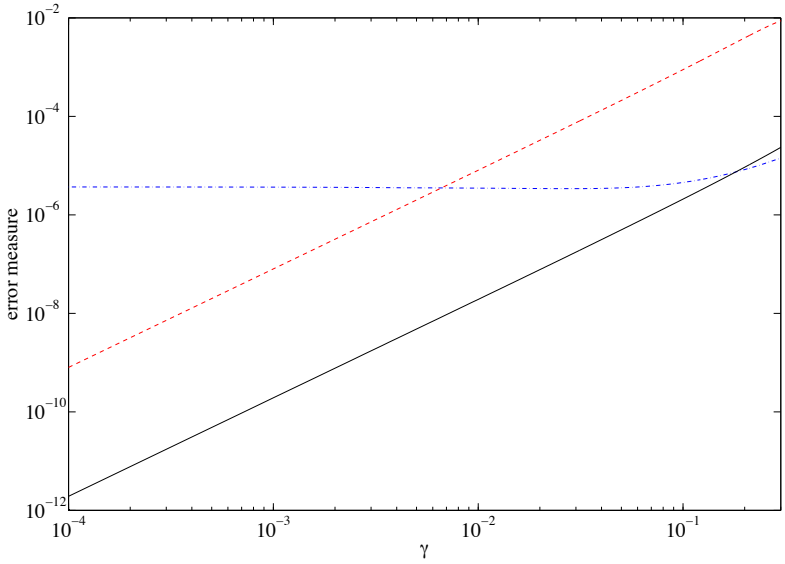


Figure 2.15: The solid line shows the maximum over time of the squared end point error in the solution, computed using the exact DNRBC without truncation, as γ is varied. The dashed line shows the maximum over time of r, the reflection ratio, as γ increases. The dot-dashed line shows the maximum squared end point error in the solution when using eight optimal coefficients.

where bars and hats respectively indicate filtered and transformed quantities. Solving for $\hat{\bar{u}}_i$ and substituting in the transformed version of (2.56) we obtain

$$\frac{\hat{u}_m}{\Delta t^2}\left[\mathrm{e}^{2\pi \mathrm{i}l/N} - 2 + \frac{\mathrm{e}^{-2\pi \mathrm{i}l/N}(1-2\gamma+\gamma\mathrm{e}^{2\pi \mathrm{i}l/N})}{1-\gamma\mathrm{e}^{-2\pi \mathrm{i}l/N}}\right] + \alpha\hat{u}_m - c^2\frac{\hat{u}_{m+1}-2\hat{u}_m+\hat{u}_{m-1}}{\Delta x^2} = 0. \quad (2.58)$$

Finally we solve this difference equation by introducing $\hat{u}_m = \bar{\kappa}^m$, where the bar is used to differentiate this quantity from that derived for the unfiltered case:

$$\frac{1}{\Delta t^2}\left[\mathrm{e}^{2\pi \mathrm{i}l/N} - 2 + \frac{e^{-2\pi \mathrm{i}l/N}(1-2\gamma+\gamma\mathrm{e}^{2\pi \mathrm{i}l/N})}{1-\gamma\mathrm{e}^{-2\pi \mathrm{i}l/N}}\right] + \alpha - c^2\frac{\bar{\kappa}-2+1/\bar{\kappa}}{\Delta x^2} = 0. \quad (2.59)$$

It remains only to solve for $\bar{\kappa}$ as a function of l and take its inverse DFT to obtain the coefficients, $\bar{\kappa}_k$. The resulting coefficients are plotted in figure 2.16 for three values of γ, with our usual choice of other parameters. The sets of coefficients computed by DFT, κ_k and $\bar{\kappa}_k$, were both applied without truncation to the test problem over 1000 time steps; both series were calculated using $N = 10^5$. The results of testing both without filtering and with strong filtering are shown in table 2.3. It is clear that the boundary condition derived to account for the filtering performs much better for the large value of γ than the original boundary condition, which does less well in this test of rather excessive filtering.

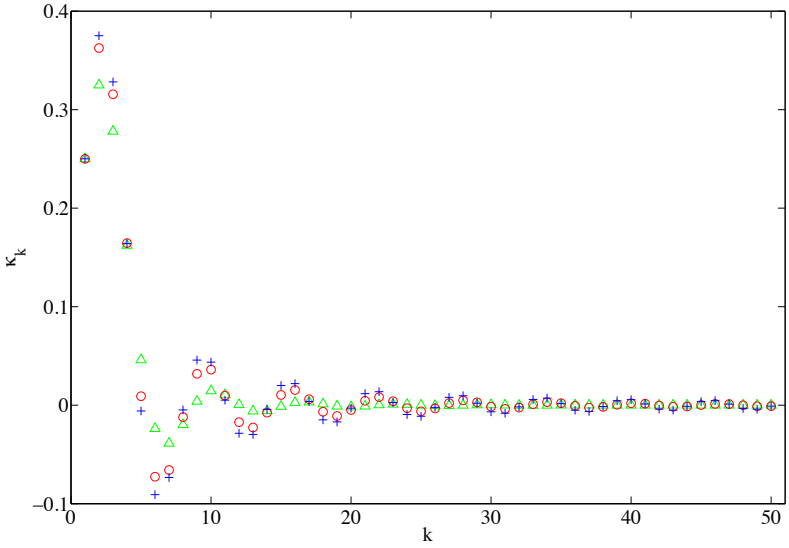


Figure 2.16: The first fifty convolution coefficients obtained by taking account of the time filtering, computed using 10^5 sampling points. Crosses: $\gamma = 0$; circles: $\gamma = 0.05$; triangles: $\gamma = 0.2$. Other parameters: $c = \alpha = 1$, $\Delta x = 0.01$, $\Delta t = 0.005$. Notice the more rapid decay rate as $k \to \infty$, when γ is increased.

γ	κ_k		$\bar{\kappa}_k$	
	max. squared end point error	$r(500\Delta t)$	max. squared end point error	$r(500\Delta t)$
0	1.8320×10^{-5}	0.0083	1.8320×10^{-5}	0.0083
0.3	1.5038×10^{-5}	0.0183	2.8854×10^{-6}	0.0015
0.5	8.6053×10^{-5}	0.0253	2.8125×10^{-6}	0.0003

Table 2.3: Results of comparison of boundary conditions that either neglect or account for time filtering. As before, $c = \alpha = 1$, $\Delta x = 0.01$, $\Delta t = 0.005$.

2.3.7 Discussion

In this section, the performance of a practical, exact, discretely non-reflecting boundary condition for the Klein-Gordon equation has been tested, and the limitations on its effectiveness when the convolution coefficients are truncated has been investigated. This is the scheme that will be implemented when solving subsequent problems of this type numerically. A stability analysis has revealed why the truncated scheme is unstable in some cases, and provided a condition useful when selecting coefficients. A method of choosing such coefficients as the solution to a minimisation problem has been presented, and the resulting values have been shown to perform favourably when used to form the boundary condition. Finally, the fact that time filtering the solution renders such DNRBCs non-exact has been explored. The increase in reflection error as the filtering is stepped up has been quantified, and a boundary condition that remedies the problem has been derived. The influence of time-filtering schemes on the performance of a non-reflecting boundary condition is an issue which has so far been overlooked in the literature, but is an important one given the relative effort invested in the development of an adequate outflow condition. In particular, the new simple stability criterion derived is a particularly useful tool since it can be added as a constraint to the optimisation procedure, in order to guarantee stability.

In this section we have studied the boundary condition of Han & Zhang, which is derived for a particular PDE with a particular choice of discretisation scheme, since they are the PDE and solution technique that are relevant for 1-d model problems in subsequent chapters. However, taking a broader view of these developments, it is hoped that the essence of the work could be applied in other settings; the Klein-Gordon equation solved using what is a rather basic finite-difference scheme, merely serves as a prototype system. It would be of interest to see if ideas explored using this simple numerical scheme could be extended and applied in more practical settings — i.e. PDE systems describing physical phenomena, solved using more sophisticated numerical schemes, such as a symplectic integrator for the nonlinear Klein-Gordon equation [17] — in order to yield improved non-reflecting boundary conditions for realistic physical models.

The results of this section — in particular the exact DNRBC accounting for Asselin filtering and the DFT method for computing the coefficients — are employed in subsequent chapters, which feature simulation of wave radiation to infinity in the context of the Klein-Gordon equation. The exact non-reflecting boundary condition is employed in the numerical schemes used in chapters 3 and 4, using the full DNRBC without truncation.

Having studied a prototype one-dimensional problem, we now move on to a 2-d problem with a practical application, by considering a simple, accurate and efficient class of boundary conditions for the rotating shallow water equations.

2.4 An effective 2-d linearised shallow water solver

Having developed useful numerical schemes for the solution of relevant one-dimensional model problems, we find ourselves in the appropriate place to consider how best to solve some of the more physically realistic equations that will be used to model wave radiation — the shallow water equations. In this section we focus on the linearised shallow water equations in two dimensions. It is desirable to find a numerical method that possesses the following features:

- a discrete dispersion relation that mirrors that of the continuous equations, e.g. with one vortical mode and two gravity wave modes;
- efficient, accurate and stable boundary non-reflecting boundary conditions on whichever boundaries one chooses;
- the possibility of solving either the linearised or fully nonlinear problem;
- an interior scheme that is efficient yet relatively simple.

Given that the full equations are nonlinear, the problem will be solved using a splitting method, by separating each of the equations into a linear and nonlinear part. Such a scheme has the benefit of allowing the nonlinear terms to be "switched off" should the need arise. With a view to satisfying the first and fourth requirements, we choose to use a staggered-time/staggered-space finite-difference scheme. By using a staggered method we gain an effective increase in numerical resolution due to the averaging properties of the scheme; we also obtain a scheme whose dispersion behaviour approximates that of the continuous problem fairly well.

The second requirement is less of a simple matter; although there has been much research in finding NRBCs for dispersive systems in more than one dimension, choosing a boundary condition is not easy. Recently, Dea [15] investigated the use of Higdon NRBCs — as mentioned in §2.3 — for solving the linearised shallow water equations solved on a staggered grid. Numerical experiments suggest that such an application is a useful one, but no stability analysis exists for this combination of interior scheme and boundary condition. Also, Higdon's NRBCs require some parameters to be chosen in order to control their performance; Dea made a simplification to the scheme which removed this freedom of parameter choice. We address both of these issues in this section.

To deal with the algebraic complexity of the high-order Higdon NRBC, we derive a novel, efficient and easily programmable algorithm for its implementation, for arbitrary order. We then evaluate the boundary condition's effectiveness by numerical experiment. Givoli & Neta proposed an algorithm for choosing the free parameters in a general Higdon NRBC, and we take the step of using the strategy to generate the parameters for the scheme in question. We compare the results using the Givoli-Neta algorithm, with those found by making a simple choice for the free parameters. The scheme has not previously undergone a stability analysis, so we use semi-numerical techniques to examine its stability, which has implications for the parameters that may be used with it. The scheme is also extended to handle the nonlinear terms in the SWEs. The main new results are that the process for choosing the coefficients suggested by Givoli and Neta does not offer any improvement in this case, and that the scheme investigated by Dea may be unstable in certain circumstances.

The outline of the remainder of this section is as follows. In §2.4.1 the splitting method for the shallow water equations is first described. In §2.4.2–2.4.4 the Higdon NRBC [39] and its discrete form are introduced, and in §2.4.5 an efficient implementation is derived. In §2.4.6–2.4.8 the Higdon NRBC is applied to the linearised shallow water equations; we first repeat the experiments of Dea, and then go further by allowing the parameters in the NRBC to be chosen using a method due to Givoli & Neta, and compare the results with those found using the simplified scheme. In §2.4.9–2.4.12 the stability of the combined interior difference scheme and NRBC is analysed, and we derive some new constraints on the model/scheme parameters necessary for stability.

2.4.1 A splitting method

Consider the general partial differential equation

$$\frac{\partial \psi}{\partial t} + \mathcal{F}(\psi) = 0, \tag{2.60}$$

where $\mathcal{F}$ is an operator independent of time, and ψ is some continuous function. By formally integrating (2.60) over an interval of time Δt, we obtain the solution at an advanced time

$$\psi(t + \Delta t) = \exp(\Delta t \mathcal{F})\psi(t). \tag{2.61}$$

The idea of operator splitting is as follows: suppose the operator $\mathcal{F}$ may be decomposed into $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2$. Now, introducing discrete approximations to these operators which we denote $\mathcal{G}_1$ and $\mathcal{G}_2$, and the approximate solution at the nth time level ϕ^n, the solution may be advanced by a two-step process:

$$\hat{\phi} = \mathcal{G}_1(\Delta t)\phi^n, \tag{2.62}$$

$$\phi^{n+1} = \mathcal{G}_2(\Delta t)\hat{\phi}. \tag{2.63}$$

Such a method is referred to as a *split-step, fractional-step,* or *splitting* method.

Consider such a technique as applied to the nonlinear shallow water equations. We have the following system

$$\frac{\partial \mathbf{s}}{\partial t} + \mathcal{F}_{\mathrm{L}}(\mathbf{s}) + \mathcal{F}_{\mathrm{NL}}(\mathbf{s}) = \mathbf{0}, \tag{2.64}$$

where

$$\mathbf{s} = \begin{pmatrix} u \\ v \\ \eta \end{pmatrix}, \quad \mathcal{F}_{\mathrm{L}} = \begin{pmatrix} 0 & -f & g\partial_x \\ f & 0 & g\partial_y \\ H\partial_x & H\partial_y & 0 \end{pmatrix}, \quad \mathcal{F}_{\mathrm{NL}} = \begin{pmatrix} u\partial_x + v\partial_y & 0 & 0 \\ 0 & u\partial_x + v\partial_y & 0 \\ 0 & 0 & u\partial_x + v\partial_y \end{pmatrix} \tag{2.65}$$

The equations have been split using operators $\mathcal{F}_{\mathrm{L}}$ and $\mathcal{F}_{\mathrm{NL}}$ to account for the linear and nonlinear terms, respectively. The class of problems which will be studied will usually involve a central core featuring a mixture of linear and nonlinear activity, surrounded by a wave region from which waves (that are linear to a very good approximation) emanate. Hence, the dynamics at a far-field boundary will, to within a small error, be linear. The idea is now this: we advance

the solution to the fully nonlinear problem using equations (2.62) and (2.63), but add an intermediate step where the **boundary conditions for the linear problem are applied**. For the nonlinear correction part, a Dirichlet boundary condition is imposed, i.e. $u = v = \eta = 0$ on ∂.

2.4.2 Higdon's higher order NRBC

From the literature on non-reflecting boundary conditions we see that there are many different approaches possible when deriving a non-reflecting boundary condition, even when the system in need of such a boundary condition appears simple. In §2.3.1 one possible approach was considered: the Klein-Gordon equation was first discretised using finite differences, and from this an NRBC was derived to eliminate all reflections in the discrete scheme, regardless of the time step and grid scale. The NRBC was nonlocal in time, and hence the boundary condition was truncated so that it could be employed in experiments with many time steps.

Rather than starting from an exact boundary operator that is nonlocal, one may instead begin with a boundary operator chosen ad hoc that is local in both space and time, and choose its form to ensure that it is both stable and as close to reflectionless as possible, by approximating the exact DNRBC for the scheme in question well. This is the approach to which we now turn our attention, and we are not the first. Working with the one-dimensional wave equation, Engquist & Majda [23] chose a form of boundary operator that was practical, whilst featuring a parameter. Their left-hand boundary condition for the value u_0^n in terms of $u_1^n, u_0^{n-1}, u_1^{n-1}, u_0^{n-2}$ and u_1^{n-2} is of the form

$$\left(\frac{D_t^+ + D_t^-}{2}\right)\left(\frac{u_0^{n-1} + u_1^{n-1}}{2}\right) - \theta D_x^+ u_0^{n-1} - (1-\theta) D_x^+ \left(\frac{u_0^n + u_0^{n-2}}{2}\right) = 0. \tag{2.66}$$

where θ is the parameter to be determined, and the operators D^+ and D^- represent respectively forward and backward differences in either time or space, e.g.

$$D_t^+ u_m^n = (u_m^{n+1} - u_m^n)/\Delta t; \quad D_x^- u_m^n = (u_m^n - u_{m-1}^n)/\Delta x. \tag{2.67}$$

An optimal choice of θ corresponds to making the discrete dispersion relation of (2.66) conform as closely as possible to that of the finite-difference scheme for the wave equation:

$$\left(D_t^+ D_t^- - D_x^+ D_x^-\right) u_m^n = 0. \tag{2.68}$$

An Engquist & Majda note, there are many possible measures of "closeness" between the two dispersion relations, leading to several possible "good" choices of θ. For example, by Taylor expanding both (2.66) and (2.68), all error terms in Δx and Δt or order four and above are cancelled by taking

$$\theta = \frac{3\lambda^2 - 1}{4\lambda^2}. \tag{2.69}$$

The problem remains that however good a choice of the parameter θ is made, the boundary condition is still only second-order in time and first-order in space, i.e. the boundary operator is of the form $\sum_{i=0}^{2}\sum_{j=0}^{2} a_{i,j} S_t^{-i} S_x^{+j} u_0^n$. Here we have used the shift operators defined as follows:

$$S_t^+ u_m^n = u_m^{n+1}, \quad S_t^- u_m^n = u_m^{n-1}, \quad S_x^+ u_m^n = u_{m+1}^n, \quad S_x^- u_m^n = u_{m-1}^n, \tag{2.70}$$

and introduced the notation

$$S_t^{+k} = (S_t^+)^k. \tag{2.71}$$

In a series of papers, Higdon [39, 40, 41, 42, 43] analysed a class of boundary condition of arbitrarily high order, henceforth referred to as Higdon NRBCs. The higher order Higdon NRBC is simply a product of multiple Sommerfeld operators, i.e. the first-order Higdon NRBC is simply the basic boundary condition derived by factorising the wave equation. Higdon originally derived such boundary conditions for the non-dispersive wave equation in [39], but in [43] was able to show that it can be applied to the Klein-Gordon equation with equal success.

Givoli and his associated collaborators have been instrumental in investigating the application of Higdon NRBCs to practical numerical problems. Givoli & Neta [29] first tested the boundary condition used at high order for the Klein-Gordon equation, and then went on to investigate its usefulness for the shallow water equations [30]. In [29] they proposed a useful algorithm to make application of the boundary condition more straightforward. Neta et al [61] used this work to study Higdon NRBCs applied to shallow water even further — work which is also contained in the thesis of Van Joolen [78]. More recently, Dea et al [16] applied Higdon NRBCs to the 2-d linearised Euler equations with rotation. Lately, Dea [15] provided several examples of systems that are compatible with a Higdon NRBC, including the linearised shallow water equations and the free space Maxwell equations. In particular, the work dealt with the shallow water equations **discretised on a staggered grid with staggering in time**, and showed that even for this system, which in its discrete form is no longer equivalent to the Klein-Gordon equation for which the boundary condition was constructed, the results are very favourable. Rather than use the full algorithm derived in [29] to implement the NRBC, a simplification was made.

In the following section, we study the linearised shallow water equations in the presence of rotation and continue the work of Dea, following lines of enquiry suggested by his conclusions to provide some new results. The aims are twofold: first, to investigate the effectiveness of the full NRBC (without the simplifying procedure), and compare the two approaches. Second, we aim to determine any conditions for stability of the scheme, beyond those necessary for stability of the interior discretisation.

2.4.3 Definition and properties of Higdon NRBCs

Consider the Klein Gordon equation

$$u_{tt} - c^2\nabla^2 u + \alpha u = 0, \tag{2.72}$$

solved on the half plane $x > 0$, with a non-reflecting boundary required on the $x = 0$ boundary. The Higdon NRBC for the continuous equation can be written in the form

$$\left[\prod_{j=1}^{J} ((\cos(\theta_j)\partial_t - c\partial_x)\right] u(x = 0, t) = 0. \tag{2.73}$$

The constants θ_j are free to be chosen, and typically $J < 10$. Such a boundary condition is natural because for the continuous equation with $\alpha = 0$ (the dispersionless wave equation), the operator $\cos(\theta)\partial_t - c\partial_x$ cancels out plane waves travelling out of the domain $x > 0$ with

wavespeed c and angle of incidence to the boundary of $\pm\theta$. For waves following other courses, there will be some nonzero reflection. Let us write the reflection coefficient of the Higdon boundary operator as R_{H}, and consider outgoing and incoming solutions of (2.72) with amplitudes A_{out} and A_{in} and angle of incidence, ϕ.

$$A_{\mathrm{out}}\mathrm{e}^{\mathrm{i}(\cos(\phi)x+\sin(\phi)y-ct)} + A_{\mathrm{in}}\mathrm{e}^{\mathrm{i}(\cos(\phi)x+\sin(\phi)y+ct)}. \tag{2.74}$$

Now, the reflection coefficient of (2.73), for waves travelling out of the positive x domain, can be shown to have the following property when applied to (2.72) with $\alpha \neq 0$:

$$R_{\mathrm{H}} = \prod_{j=1}^{J} \left| \frac{\cos(\theta_j) - \cos(\phi)}{\cos(\theta_j) + \cos(\phi)} \right|, \tag{2.75}$$

where ϕ is the angle of incidence of the waves. The angles θ_j may be chosen to give a range of directions for which the boundary condition is perfectly non-reflecting for such waves, by perhaps using knowledge of the wave field. Notice that even if the choices of θ_j are poorly informed, because each factor in (2.75) is < 1 for $-\pi/2 < \phi < \pi/2$ (corresponding to waves leaving the positive x domain), $R_{\mathrm{H}} \to 0$ as $J \to \infty$.

Instead of considering a range of angles for which waves should pass unreflected, the boundary condition (2.73) can instead be written as

$$\left[\prod_{j=1}^{m} (\partial_t - c_j \partial_x) \right] u(x = 0, t) = 0\,; \tag{2.76}$$

here c_j is a phase velocity in the x-direction, and now the c_js denote phase velocities of waves for which the boundary condition is non-reflecting. This is easily shown for the non-dispersive situation by taking a wave of the form $A\exp(i(kx+ly-\omega t+\psi))$, and substituting in (2.73) with $\alpha = 0$; when the phase velocity in the x-direction $c_x = c_j$ the boundary condition is satisfied exactly and the wave is transmitted perfectly. It is this formulation with which we will work. Again we must choose our c_js, but can be assured that their values are not critical. Let us consider the reflection coefficient for $\alpha \neq 0$: if we take a linear superposition of incoming and outgoing waves at the boundary of the form

$$a_{\mathrm{in}}\mathrm{e}^{\mathrm{i}(k_{\mathrm{in}}x+ly-\omega t)} + a_{\mathrm{out}}\mathrm{e}^{\mathrm{i}(k_{\mathrm{out}}x+ly-\omega t)}, \tag{2.77}$$

and let $k_{\mathrm{in}} = k$ and set $k_{\mathrm{out}} = -k_{\mathrm{in}} = -k$. Then

$$R_{\mathrm{H}} = \left| \frac{a_{\mathrm{in}}}{a_{\mathrm{out}}} \right| = \prod_{j=1}^{J} \left| \frac{-\omega/c_j + k}{-\omega/c_j - k} \right|, \tag{2.78}$$

and since from (2.72) $c \to \omega/k$ in the limit of $k \to \infty$,

$$\lim_{|k|\to\infty} R_{\mathrm{H}} = \prod_{j=1}^{J} \left| \frac{c - c_j}{c + c_j} \right|, \tag{2.79}$$

where again each factor is less than unity. Expanding (2.76) we find its general form

$$\sum_{j=0}^{J} \alpha_j \partial_x^j \partial_t^{j-J} u = 0. \tag{2.80}$$

The boundary operator is symmetric, i.e. it involves partial derivatives in both space and time up to order J.

It is clear that the Higdon NRBC prior to discretisation has several features of merit:

- it is simple in form, being a product of a simple operator. This makes it relatively straightforward to analyse;
- it is local in space and time, and the order of the NRBC may be chosen freely;
- it caters for dispersive systems, and in more than one space dimension;
- the reflection coefficient is less than unity, regardless of the choices of c_j.

These strengths notwithstanding, there are some potential drawbacks. When J is large the NRBC is a product of considerable algebraic complexity. Also, the sub-unity property of the reflection coefficient is not a particularly strong one: it may be that J needs to be taken to be very large to give good results. It is also remains necessary to develop some rationale for choosing good values for c_j. We now proceed to overcome some of these obstacles to applying the Higdon boundary condition.

2.4.4 Discrete approximations to Higdon's NRBC

We discretise (2.76) using standard first order differences. Take for example the solution variable $\phi_m^n = \phi(m\Delta x, n\Delta t)$, $x \in [0, M\Delta t]$. Then the discrete Higdon NRBC of order J for the solution right-hand boundary point for which $m = M$ is

$$\begin{aligned} H^J \phi_M^n &= \prod_{j=1}^{J} \left(D_t^- + c_j D_x^-\right) \phi_M^n && (2.81)\\ &= \prod_{j=1}^{J} \left(\frac{I - S_t^-}{\Delta t} + c_j \frac{I - S_x^-}{\Delta x}\right) \phi_M^n && (2.82)\\ &= \sum_{\delta=0}^{J} \sum_{\beta=0}^{J-\delta} A_{\delta,\beta}\, S_x^{-\delta} S_t^{-\beta} \phi_M^n && (2.83)\\ &= 0, \end{aligned}$$

where I is the identity operator. The discretised NRBC gives an expression for the boundary solution at the nth time level, in terms of its previous history ($\phi_M^{n-1}, \phi_M^{n-2}$, etc.), and the interior solution at several time levels up to and including the nth one. As such, the scheme gives an explicit update rule for the boundary. The boundary condition is symmetric, i.e. each term comprises at most J shifts in space or time. The NRBC involves solution information from points close to the boundary further back in time than those points that lie away from the boundary. Expanding (2.83) we see that that the solution is required at a total of $(J+1)(J+2)/2 - 1$ points in space and time; figure 2.17 illustrates this.

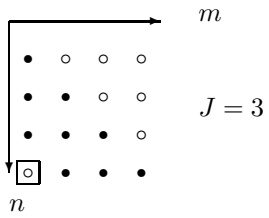


Figure 2.17: The updated boundary value's dependence on the solution at points in time and space. The boundary value marked with a box is to be updated, using the solution at points marked with filled circles. Note that roughly half the stored solution is unused.

In [43], it was shown that the boundary operator discretised in this way is GKS stable, provided that the interior difference scheme is the centred second-order one, i.e.

$$D_t^+ D_t^-(u_m^n) - c^2 D_x^+ D_x^-(u_m^n) + \alpha u_m^n = 0. \tag{2.84}$$

Since the NRBC is derived for the continuous equation, we must expect that it will not be applicable to problems solved using a coarse grid or large time step, unlike a true DNRBC. Higdon [43] and subsequently Givoli & Neta [29] also considered a more sophisticated alternative to the basic discretisation in (2.81), by combining the approximation to the derivatives using weighted averages, giving the following boundary condition

$$\prod_{j=1}^{J}\left[\frac{I-S_t^-}{\Delta t}((1-b)I+bS_x^-) + c_j\frac{I-S_x^-}{\Delta x}((1-b)I+bS_t^-)\right]\phi_M^n = 0. \tag{2.85}$$

For simplicity, we will consider only the case of $b = 0$.

With a little thought, the NRBC may also be applied to problems in 2 dimensions. Consider a grid with x-index i and y-index j. We retain n as the time index, and write a solution as $\phi_{i,j}^n = \phi(i\Delta x, j\Delta y, n\Delta t)$. We refer to the boundary furthest in the x-direction as the South boundary, and that furthest in the y-direction as the East boundary. This sets our coordinate system with the origin at the top left corner. Figure 2.18 makes this clear. We employ Dea's

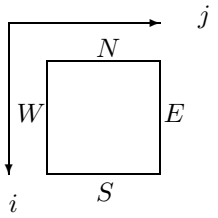


Figure 2.18: Coordinate axes and boundary notation for two-dimensional problems.

notation by using N, E, S and W as indices on boundaries; for example, $\phi_{i,E}$ is the ith point in the x-direction on the Eastern boundary. It is now a straightforward matter to write down

the NRBCs applied to fields at all four boundaries:

$$\text{North: } \prod_{j=1}^{J}\left(\frac{I-S_t^-}{\Delta t}+c_j\frac{I-S_x^+}{\Delta x}\right)\phi_{N,j}^n=0; \tag{2.86a}$$

$$\text{East: } \prod_{j=1}^{J}\left(\frac{I-S_t^-}{\Delta t}+c_j\frac{I-S_x^-}{\Delta x}\right)\phi_{i,E}^n=0; \tag{2.86b}$$

$$\text{South: } \prod_{j=1}^{J}\left(\frac{I-S_t^-}{\Delta t}+c_j\frac{I-S_x^-}{\Delta x}\right)\phi_{S,j}^n=0; \tag{2.86c}$$

$$\text{West: } \prod_{j=1}^{J}\left(\frac{I-S_t^-}{\Delta t}+c_j\frac{I-S_x^+}{\Delta x}\right)\phi_{i,W}^n=0. \tag{2.86d}$$

2.4.5 Implementation issues

For the one-dimensional problem with $J=1$, equation (2.81) is easily solved for ϕ_M^n, yielding:

$$\phi_M^n=\left(\frac{\phi_M^{n-1}}{\Delta t}+c_1\frac{\phi_{M-1}^n}{\Delta x}\right)\Big/\left(\frac{1}{\Delta t}+\frac{c1}{\Delta x}\right). \tag{2.87}$$

and for $J=2$

$$\phi_M^n=\left[\frac{c_1c_2}{\Delta t\Delta x}\phi_M^{n-1}+\frac{2}{\Delta t^2}\phi_M^{n-1}+\left(\frac{c_2}{\Delta t\Delta x}+\frac{c_1}{\Delta t\Delta x}+\frac{2c_1c_2}{\Delta x^2}\right)\phi_{M-1}^n \tag{2.88}$$

$$+\left(\frac{-c_2}{\Delta t\Delta x}-\frac{c_1}{\Delta t\Delta x}\right)\phi_{M-1}^{n-1}-\frac{c_1c_2}{\Delta x^2}\phi_{M-2}^n-\frac{1}{\Delta t^2}\phi_M^{n-2}\Bigg] \tag{2.89}$$

$$\Big/\left[\frac{1}{\Delta t^2}+\frac{c_2}{\Delta t\Delta x}+\frac{c_1}{\Delta t\Delta x}+\frac{c_1c_2}{\Delta x^2}\right]. \tag{2.90}$$

At this point it is evident that implementation of the Higdon NRBC may require a little thought. Clearly, computing the coefficients $A_{i,j}$ by hand will be rather involved for $J>3$ (indeed, $J=3$ is as far as Higdon pursued the calculation), and so an automatic method must be derived or a simplification made. Dea, working with the rotating SWEs, chose to make a simplification that had been adopted in [61]: by choosing $c_j\equiv c$, where c is the wavespeed in the equation (i.e. $\sqrt{gH}$ for the shallow water equations), the coefficients $A_{\delta,\beta}$ in (2.83) have a simple general form, and the boundary condition reduces from a product to a sum

$$H^J\phi_M^n=\sum_{\delta=0}^{J}\sum_{\beta=0}^{J-\delta}a^\delta b^\beta d^\gamma S_x^{-\gamma}S_t^{-\beta}\phi_m^n=0, \tag{2.91}$$

where

$$\gamma=J-\delta-\beta \tag{2.92}$$

$$a=1-c \tag{2.93}$$

$$b=-1 \tag{2.94}$$

$$d=-c\Delta t/\Delta x \tag{2.95}$$

However, by making this simplification, we lose the freedom to tune the coefficients c_j, and so we do not pursue this idea, and instead work with the NRBC in its full generality.

A method for implementing the boundary condition can be constructed that is both intuitive and simple to program, and which produces the coefficients $A_{\delta,\beta}$ in H^J. We choose to store our solution in memory using a column vector with the following structure:

$$\mathbf{u}^n = [\overbrace{u_M^n, u_{M-1}^n, u_{M-2}^n, \ldots, u_{M-J}^n}^{\text{n}}, \overbrace{u_M^{n-1}, \ldots, u_{M-J}^{n-1}}^{n-1}, \ldots, \overbrace{u_M^{n-J}, \ldots, u_{M-J}^{n_J}}^{n-J}]^T. \tag{2.96}$$

$J+1$ total sets of $J+1$ terms

This is not an unnatural choice, since a memory update is easily achieved by shifting the entries up by M places. Having chosen this arrangement for the solution, we seek a boundary update row vector $\mathbf{z}^J$ such that

$$\mathbf{z}^J \mathbf{u}^n = u_M^{n+1}. \tag{2.97}$$

Now, in the notation of [29], multiply through (2.81) by Δt to write the NRBC as

$$H^J \equiv \left[\prod_{j=1}^{J} (a_j I + d\, S_t^- + e_j S_x^-) \right] u_M^n = 0. \tag{2.98}$$

Here we have defined

$$a_j = 1 + \frac{c_j \Delta t}{\Delta x}, \tag{2.99}$$

$$d = -1, \tag{2.100}$$

$$e_j = -\frac{c_j \Delta t}{\Delta x}. \tag{2.101}$$

As mentioned earlier, Givoli & Neta [29] proposed one algorithm for computing the coefficients in the expansion of (2.81), which is as follows. Write the Higdon boundary operator as

$$H^J = \sum_{m=0}^{3^J-1} A_m P_m u_M^n = 0. \tag{2.102}$$

Solving for u_M^n (the updated value at the boundary) gives

$$u_M^n = -\frac{Z^*}{A_0}, \tag{2.103}$$

where

$$Z^* = \sum_{m=1}^{3^J-1} A_m P_m u_M^n, \tag{2.104}$$

with A_m being a coefficient and P_m a product of the operators S_x, S_t and I. The algorithm (given in [29]) for determining Z^* is as follows.

Algorithm 1 Givoli-Neta algorithm for implementing H^J

$Z^* := 0$, $A_0 := \prod_{j=1}^{J} a_j$
for $m := 1, \ldots, 3^J - 1$ **do**
  Write the number m in base 3 using J digits, and store the digits in the vector $D_r(j)$, $j = 1, \ldots, J$
  $A_m := 1$
  for $j = 1, \ldots, J$ **do**
    if $D_r(j) = 0$ **then**
      $A_m \leftarrow a_j A_m$
    else if $D_r(j) = 1$ **then**
      $A_m \leftarrow d_j A_m$
    else
      $A_m \leftarrow e_j A_m$
    end if
  end for
  $P_m := u_M^n$
  for j = 1,. . . ,J **do**
    if $D_r(j) = 1$ **then**
      $P_m \leftarrow S_t^{-1} P_m$
    else if $D_r(j) = 2$ **then**
      $P_m \leftarrow S_x^{-1} P_m$
    end if
  end for
  $Z^* \leftarrow Z^* + A_m P_m$
end for
$u_M^n = -\frac{Z^*}{A_0}$

Since H^J is simply a product of discrete operators, with each involving single shifts in time and space, we may construct H^J and hence $\mathbf{z}^J$ *recursively*. We now derive a new algorithm for determining the form of the boundary condition conveniently and efficiently. The resulting algorithm avoids the costly loop over 3^J terms involved in the Givoli-Neta algorithm. Our algorithm is as follows.

Algorithm 2 Recursive construction of $\mathbf{z}^J$

$\mathbf{a}_1 := [a_1, d, e_1]$
$\mathbf{a} := \mathbf{a}_1$
$\mathbf{M} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$
$j := 1$
while $j < J$ **do**
  $j := j + 1$
  $\mathbf{a} \leftarrow \mathbf{a}_1 \otimes \mathbf{a}$
  for $k := 1, \ldots, j$ **do**
    insert column of zeros after j columns of $\mathbf{M}$ to form $\mathbf{M}^*$.
  end for
  $\mathbf{M} \leftarrow \begin{bmatrix} \mathbf{M}^* \\ \mathbf{M}^* \\ \mathbf{M}^* \end{bmatrix}$
  shift entries in middle third of $\mathbf{M}$ right by $j + 1$ places
  shift entries in middle third of $\mathbf{M}$ right by one place
end while
$A_0 := \text{prod}(\mathbf{a})$
$\mathbf{z}^J := -\mathbf{a}\mathbf{M}/A_0$
set first entry of $\mathbf{z}^J$ to zero

Here we have used $\otimes$ to represent the matrix Kronecker product, defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}, \tag{2.105}$$

where $\mathbf{A}$ is an $m \times n$ matrix. The function prod$(\cdot)$ is defined as the product of all the elements of a vector. To understand the derivation of the algorithm, consider H^1 which corresponds to the vector $\mathbf{z}$ which multiplies the vector

$$\mathbf{u}^n = [u_M^n, u_{M-1}^n, u_M^{n-1}, u_{M-1}^{n-1}]^T. \tag{2.106}$$

Each row of M generates a term corresponding to an operator in the expanded form of $\mathbf{z}^1$; the first row corresponds to Iu_M^n which is the first entry of $\mathbf{u}^n$, hence there is a 1 in the first column of this row, the second corresponds to $S_t^- u_M^n$ which is the third entry of $\mathbf{z}^1$, hence there is a 1 in the third column of this row, and similarly for the third row. The coefficients are handled by the vector $\mathbf{a}$. At each step the process simply involves multiplying by another factor of the product (the Kronecker product step), and adjusting $\mathbf{M}$ to handle the new version of $\mathbf{u}^n$ that includes data from more time/space levels (done by the copying and shifting procedures). In effect nothing more is required than some careful bookkeeping.

Algorithm 2 is conveniently implemented in a few lines of MATLAB. Despite the matrix $\mathbf{M}$ being large (it has $3^J(J+1)^2$ entries), it is sparse, and presents no memory problems up to $J = 13$, which is further than will be considered in this work, and also fast enough for

most purposes; the result for $J = 10$ requires 0.2s to compute[2]. Table 2.4 lists the resulting coefficients $A_{\delta,\beta}$ in the expansion of the boundary operator given by (2.83) for some typical parameter values, as far as $J = 5$.

2.4.6 Higdon NRBCs applied to the linearised shallow water equations

Let η, u and v be respectively the height, x-velocity and y-velocity perturbation fields, f the Coriolis parameter, and H the rest height, so that $h_{\text{tot}} = H + \eta$. Setting the gravitational acceleration $g = 1$, we write the linearised shallow water equations (SWEs):

$$\partial_t u + \partial_x \eta - fv = 0, \tag{2.107}$$

$$\partial_t v + \partial_y \eta + fu = 0, \tag{2.108}$$

$$\partial_t \eta + H(\partial_x u + \partial_y v) = 0. \tag{2.109}$$

It is easy to show that the linearised SWEs can be reduced to a single equation for the height perturbation

$$\partial_t(\eta_{tt} - H(\eta_{xx} + \eta_{yy}) + f^2\eta) = 0. \tag{2.110}$$

Ignoring the vortical mode corresponding to the outer time derivative, we recognise the Klein-Gordon equation. Since the Higdon NRBC is known to behave well — it is both stable and increasingly accurate as $J \to \infty$ — for (2.110) solved using a centred second-order finite-difference scheme, it is only necessary to show that the discretised versions of (2.107)–(2.109) reduce to the same finite-difference scheme for η, as done in [16] for the following discretisation:

$$\Delta_t u + \Delta_x \eta - fv = 0, \tag{2.111}$$

$$\Delta_t v + \Delta_y \eta + fu = 0, \tag{2.112}$$

$$\Delta_t \eta + H(\Delta_x u + \Delta_y v) = 0, \tag{2.113}$$

where the centred difference operator Δ_a is defined as follows

$$\Delta_a = \frac{S_a^+ - S_a^-}{\Delta a}. \tag{2.114}$$

Instead, Dea [15] chose to consider the linearised SWEs solved on a *staggered grid with staggering in time*. Although in this case the discretised versions of (2.107–2.109) do not reduce to the appropriate finite-difference scheme for η, nonetheless there is empirical evidence that the simplified NRBC with $c_j = c\ \forall j$ still performs well. We first outline the staggered scheme and then explain why it is not consistent with the second-order centred scheme for η.

The problem is solved on an Arakawa "B" grid, with the height field offset from the two velocity fields by a half grid step in both the x- and y-directions. The staggered grid is illustrated schematically in figure 2.19. The solutions are staggered in time as well as space, with the height perturbation fields evaluated at integer time levels, and velocity fields at the half steps

[2]In this thesis, all computation times (unless otherwise stated) are for a desktop DELL Optiplex 745 with an Intel® Core™2 2.4GHz processor.

	β	0					1					2					3					4					5
J	δ	1	2	3	4	5	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0
1		0.2000	-	-	-	-	0.8000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2		0.4000	−0.0400	-	-	-	1.6000	−0.3200	-	-	-	−0.6400	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3		0.6000	−0.1200	0.0080	-	-	2.4000	−0.9600	0.0960	-	-	−1.9200	0.3840	-	-	-	0.5120	-	-	-	-	-	-	-	-	-	-
4		0.8000	−0.2400	0.0320	−0.0016	-	3.2000	−1.9200	0.3840	−0.0256	-	−3.8400	1.5360	−0.1536	-	-	2.0480	−0.4096	-	-	-	−0.4096	-	-	-	-	-
5		1.0000	−0.4000	0.0800	−0.0080	0.0003	4.0000	−3.2000	0.9600	−0.1280	0.0064	−6.4000	3.8400	−0.7680	0.0512	-	5.1200	−2.0480	0.2048	-	−2.0480	0.4096	-	0.3277	-	-	-

Table 2.4: Coefficients $A_{\delta,\beta}$ of the operator $S_x^{-\delta} S_t^{-\beta}$ in the expansion of the Higdon NRBC, with $c_j = 1\ \forall j$, up to fifth order. Here $\Delta x = 0.1$, and $\Delta t = 0.025$.

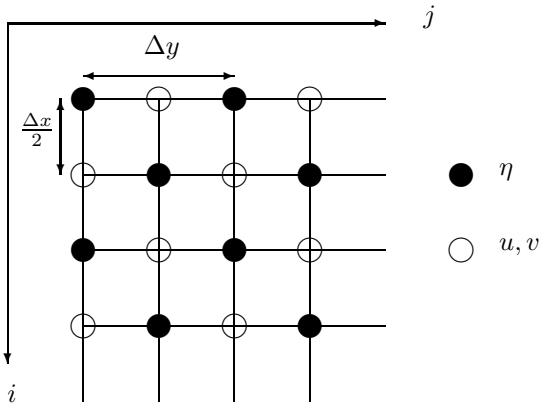


Figure 2.19: The Arakawa "B" grid used for solving the linearised SWEs.

in between. The finite-difference implementation of (2.107)–(2.109) is then

$$\Delta_t u^n_{i+\frac{1}{2},j} + \Delta_x \eta^n_{i+\frac{1}{2},j} - f v^{n-\frac{1}{2}}_{i+\frac{1}{2},j} = 0, \tag{2.115a}$$

$$\Delta_t v^n_{i+\frac{1}{2},j} + \Delta_y \eta^n_{i+\frac{1}{2},j} + f u^{n-\frac{1}{2}}_{i+\frac{1}{2},j} = 0, \tag{2.115b}$$

$$\Delta_t \eta^{n+\frac{1}{2}}_{i,j} + H\left(\Delta_x u^{n+\frac{1}{2}}_{i,j} + \Delta_y v^{n+\frac{1}{2}}_{i,j}\right) = 0. \tag{2.115c}$$

The scheme is explicit: the velocity fields at the $(n+\frac{1}{2})$th time level are first updated using η^n, $u^{n-\frac{1}{2}}$ and $v^{n-\frac{1}{2}}$, and then the height field at the $(n+1)$th time level is updated using η^n, $u^{n+\frac{1}{2}}$ and $v^{n+\frac{1}{2}}$. Together with the NRBCs for the four boundaries discretised using (2.86a)–(2.86d) and algorithm 2 for their implementation, equations (2.115a)–(2.115c) give a complete explicit staggered time-stepping scheme for the SWEs with non-reflecting boundaries.

Algorithm 3 Staggered time and space scheme

for tstep:=1:steps **do**
  update interior values of $u^{n+\frac{1}{2}}, v^{n+\frac{1}{2}}$
  apply NRBC to boundary values of $u^{n+\frac{1}{2}}, v^{n+\frac{1}{2}}$
  update interior values of η^{n+1}
  apply NRBC to boundary values of η^{n+1}
end for

The CFL condition for such a scheme can be derived quite easily by assuming a solution to (2.115a)–(2.115c) of the form $[u_0, v_0, \eta_0]\exp(i(k\Delta x + l\Delta y - \omega\Delta t))$. Doing so yields the discrete dispersion relation

$$\frac{\sin(\omega\Delta t/2)}{\Delta t} = \sqrt{H\left(\frac{\sin(k\Delta x/2)}{\Delta x^2} + \frac{\sin(l\Delta y/2)}{\Delta y^2}\right) - f^2}, \tag{2.116}$$

from which we derive the stability condition

$$\Delta t \leq \left(\frac{H}{\Delta x^2} + \frac{H}{\Delta y^2} - f^2\right)^{-\frac{1}{2}}. \tag{2.117}$$

A question remains: why can the system (2.115a)–(2.115c) not be reduced to the centred second-order scheme for the Klein-Gordon equation? The problem lies with the terms involving f: **only when the terms multiplied by f are evaluated at the nth time level** is the discrete system reducible to the Klein-Gordon equation discretised using centred, second-order differencing. Unfortunately as Dea points out [15], in the scheme we study the velocity field is evaluated at noninteger time levels, so we cannot rely on the analyses based on the equivalence of the two schemes. Indeed given that the scheme is not entirely equivalent to that for which the Higdon NRBC was derived, there is nothing to suggest that it should work well; Dea's experimental approach did however confirm that the NRBC is useful in this case.

2.4.7 Methods of determining the coefficients

Dea evaluated the NRBC in question by imposing the choice $c_j = c\,\forall j$. Recall that every factor in (2.75) is less than unity, and hence $R_\mathrm{H} \to 0$ as $J \to \infty$, regardless of the values c_j. However we cannot guarantee that the factors will be very much smaller than one with such a basic choice for the coefficients; for a low order boundary conditions, a good choice of wavespeeds c_j can reasonably be assumed to bring about an improvement in its reflection properties.

Givoli & Neta suggested one methodology for choosing the c_js. Their algorithm estimates the maximum resolvable wavenumber in the x-direction, assuming ten grid points per wavelength, to be $k_{x\max} = \pi/5\Delta x$, and similarly $k_{y\max} = \pi/5\Delta y$. Subsequently $J-1$ points are generated in the range $(0, k_{x\max})$ using a symmetric minimax formula. Given these values k_j, $k_{x\max}$ and $k_{y\max}$, the continuous dispersion relation can be used to calculate a set of frequencies ω_j and hence a set of phase speeds $c_j = \omega_j/k_j > c$; these, together with the gravest phasespeed c, are the coefficients. As an example, for the parameter choices $\Delta x = 0.1$, $\Delta t = 0.025$ (also used to find the values in table 2.4), $f = 0.1$ and $J = 5$, we find $c_1 = 1, c_2 = 1.4282, c_3 = 1.5642, c_4 = 2.0593$, and $c_5 = 5.2231$. Table 2.5 lists the boundary operator coefficients ($A_{\delta,\beta}$) obtained using these values of c_j.

In the next section, both strategies for choosing the wavespeeds are compared using numerical experiments.

2.4.8 Numerical examples

The linearised SWEs were solved on the Arakawa "B" staggered grid using the scheme as described in algorithm 3 and equations (2.115a)–(2.115c), in order to assess the effectiveness of the simplified Higdon NRBC (as was done by Dea), and then that of the full scheme, and make a comparison between the two. The grid was chosen to have 151 points in both the x- and y-directions, i.e. the grid represents a square of sides $75\Delta x$ and $75\Delta y$. For experiment 1, the other parameters are listed in table 2.6. The chosen initial conditions were $u_0 = v_0 = 0$, together with a centred Gaussian mound in the height perturbation field:

$$\eta_0(x, y) = \epsilon e^{-10((x-x_c)^2+(y-y_c)^2)}, \tag{2.118}$$

where x_c and y_c are respectively the x- and y-coordinates of the centre of the domain, and ϵ was chosen to be 0.1. The choice of 155 time steps was made by considering a wavefront in the

J	δ \ β	0					1					2				3			4		5
		1	2	3	4	5	0	1	2	3	4	0	1	2	3	0	1	2	0	1	0
1		0.2000	-	-	-	-	0.8000	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2		0.5022	−0.0604	-	-	-	1.4978	−0.3813	-	-	-	−0.5583	-	-	-	-	-	-	-	-	-
3		0.8808	−0.2470	0.0222	-	-	2.1192	−1.2677	0.1805	-	-	−1.4854	0.4534	-	-	0.3440	-	-	-	-	-
4		1.2662	−0.5763	0.1125	−0.0080	-	2.7338	−2.6461	0.8149	−0.0806	-	−2.7776	1.8312	−0.2866	-	1.2413	−0.4193	-	−0.2055	-	-
5		1.6503	−1.0496	0.3236	−0.0486	0.0028	3.3497	−4.5021	2.1778	−0.4530	0.0343	−4.4483	4.5754	−1.4983	0.1578	2.9232	−2.0514	0.3417	−0.9487	0.3420	0.1213

Table 2.5: Coefficients $A_{\delta,\beta}$ of the operator $S_x^{-\delta} S_t^{-\beta}$ determined using the Givoli Neta algorithm. Again $\Delta x = 0.1$, $\Delta t = 0.025$, and also $f = 0.01$.

54

$H = 1$
$f = 1 \times 10^{-2}$
$\Delta x = 0.1$
$\Delta y = 0.1$
$\Delta t = 0.5\Delta t_{\mathrm{CFL}} = 0.0354$
$T_{\mathrm{s}} = 155$

Table 2.6: Parameter values for experiment 1. Here Δt_{CFL} denotes the limiting time step satisfying the CFL condition, and T_{s} is the total number of time steps.

J	E_h	E_u	E_v	compute time (s)
1	8.151605×10^{-1}	9.330289×10^{-1}	9.330289×10^{-1}	90.7
2	1.563598×10^{-1}	1.778241×10^{-1}	1.778241×10^{-1}	90.5
3	3.645462×10^{-2}	4.772196×10^{-2}	4.772196×10^{-2}	92.0
4	1.068522×10^{-2}	1.388561×10^{-2}	1.388561×10^{-2}	92.2
5	3.492157×10^{-3}	4.346525×10^{-3}	4.346525×10^{-3}	94.8
6	1.204431×10^{-3}	1.614104×10^{-3}	1.614105×10^{-3}	101.0
7	4.472048×10^{-4}	6.429833×10^{-4}	6.429970×10^{-4}	113.3
8	1.860894×10^{-4}	2.816925×10^{-4}	2.815943×10^{-4}	151.0
9	1.057849×10^{-4}	1.584823×10^{-4}	1.592793×10^{-4}	265.3
10	7.093102×10^{-4}	9.755676×10^{-4}	1.148639×10^{-3}	604.8

Table 2.7: Error norms for experiment 1: an initially stationary mound of fluid, with $c_j = 1$ and $f = 1 \times 10^{-2}$.

non-dispersive case; after this time, the wavefront generated by the collapse of the mound has had sufficient time to be reflected by the walls and return to roughly the centre of the grid. This choice should ensure that the test is stringent enough: by choosing the number of time steps to be very much shorter or longer than this will not fully test the NRBC, either because waves have not yet reached it, or in the latter case because the fields are approaching quiescence.

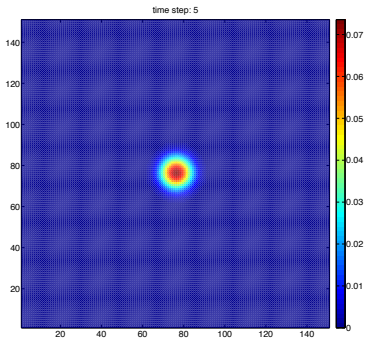
Figure 2.20 shows the boundary condition performing qualitatively well even for orders as low as $J = 3$: the height anomaly is plotted for several times in the evolution, and it is clear that it passes through the boundary without producing significant reflection.

All three fields were compared to a reference solution generated using the same scheme with a hard wall condition, and a 501×501 grid, i.e. a domain large enough that no reflections were incurred during the run. Error norms for each state variable were computed using the following formula, as used by Dea:

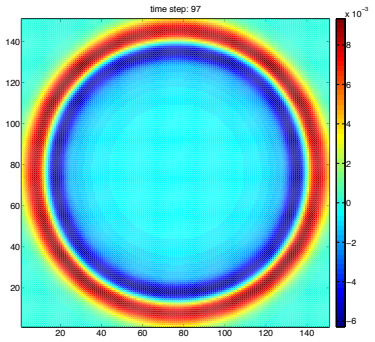
$$E_\phi = \frac{\sqrt{\sum_i \sum_j (\phi(i,j) - \phi_{\mathrm{ref}}(i,j)^2}}{\sqrt{\sum_i \sum_j \phi_{\mathrm{ref}}(i,j)^2}}. \tag{2.119}$$

Here the indices are assumed to run over the appropriate range corresponding to either the height or velocity fields. The error norms for all three fields for $1 \leq J \leq 10$, and for the the choices $f = 0.01$ and $c_j = 1$ are given in table 2.7, together with running times. Figure 2.21 shows plots of the height field error for each value of J.

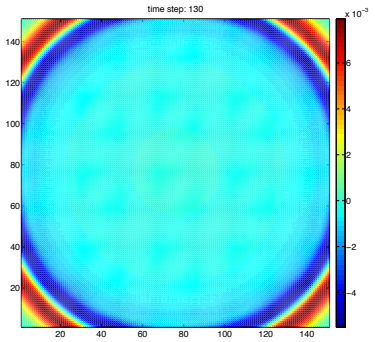
The table shows a clear drop over orders of magnitude of all three error measures, as the order J is increased. The exception is for $J = 10$, where errors are greater than for $J = 9$; a possible explanation is given later. However, the ninth order scheme is unquestionably effective, though given the large increase in running time from eighth order, it is perhaps more practical



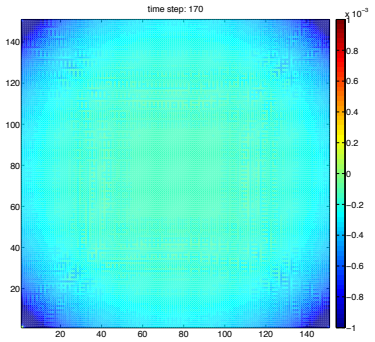
(a)



(b)

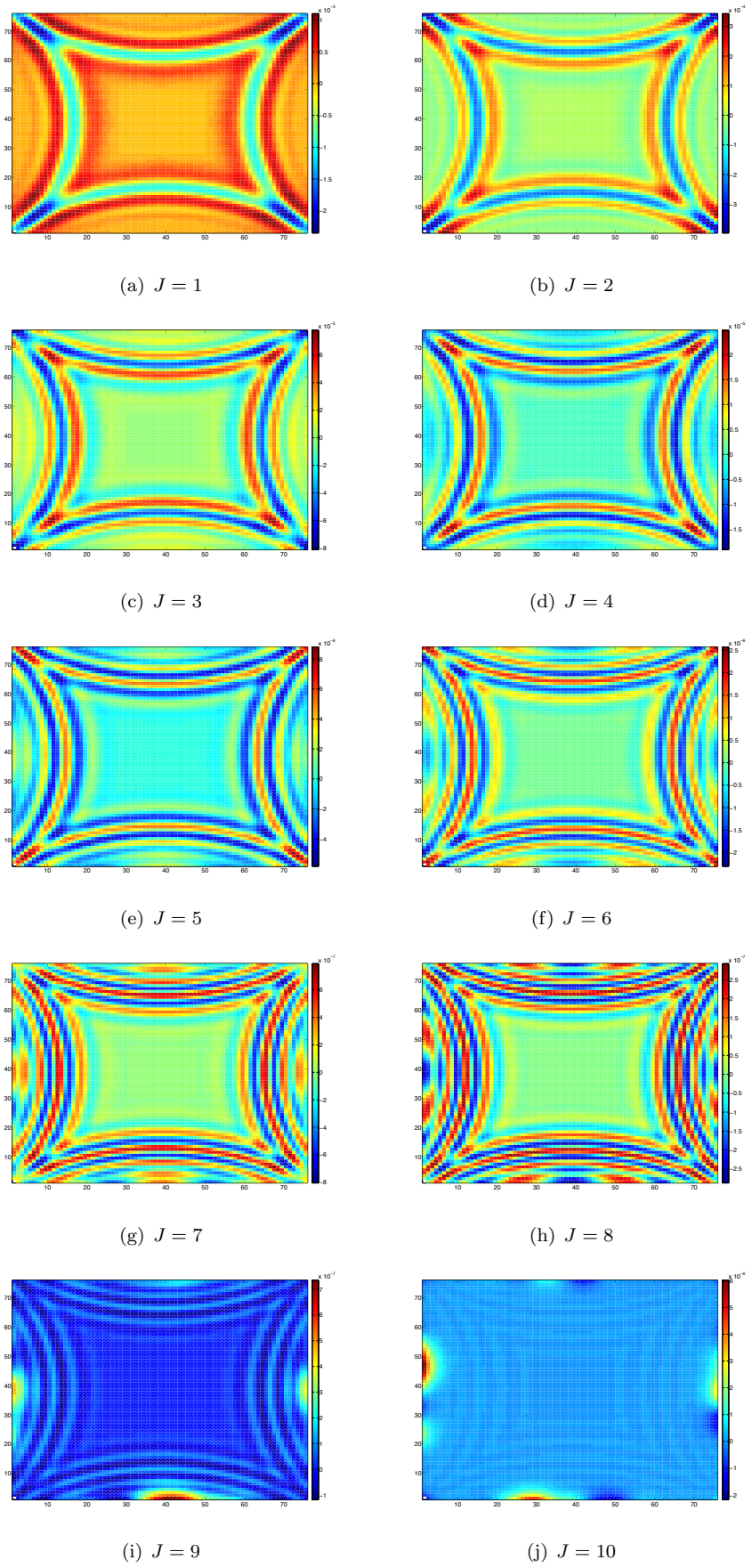


(c)



(d)

Figure 2.20: η field for the Higdon boundary condition with $J = 3$, for the test problem of the collapsing mound of fluid. There is little reflection of any significance visible in the final panel.



(a) $J = 1$ (b) $J = 2$

(c) $J = 3$ (d) $J = 4$

(e) $J = 5$ (f) $J = 6$

(g) $J = 7$ (h) $J = 8$

(i) $J = 9$ (j) $J = 10$

Figure 2.21: Error between computed and reference height fields at the final time step, for several orders of the boundary condition, with $c_j = 1 \ \forall j$.

J	E_h	E_u	E_v	compute time (s)
1	5.505955×10^{-1}	4.634918×10^{-1}	4.634918×10^{-1}	90.2
2	1.085978×10^{-1}	8.442165×10^{-2}	8.442165×10^{-2}	90.3
3	2.559083×10^{-2}	2.240105×10^{-2}	2.240105×10^{-2}	91.1
4	7.387969×10^{-3}	6.576839×10^{-3}	6.576839×10^{-3}	91.9
5	2.424007×10^{-3}	2.059626×10^{-3}	2.059626×10^{-3}	92.7
6	8.399129×10^{-4}	7.615345×10^{-4}	7.615347×10^{-4}	97.4
7	3.108402×10^{-4}	3.043899×10^{-4}	3.043869×10^{-4}	109.1
8	1.293362×10^{-4}	1.332192×10^{-4}	1.332134×10^{-4}	146.0
9	5.947607×10^{-5}	6.497576×10^{-5}	6.675863×10^{-5}	258.6
10	4.179681×10^{-4}	3.943859×10^{-4}	4.568135×10^{-4}	592.3

Table 2.8: Error norms for experiment 2: as in experiment 1, but with $f = 0.5$.

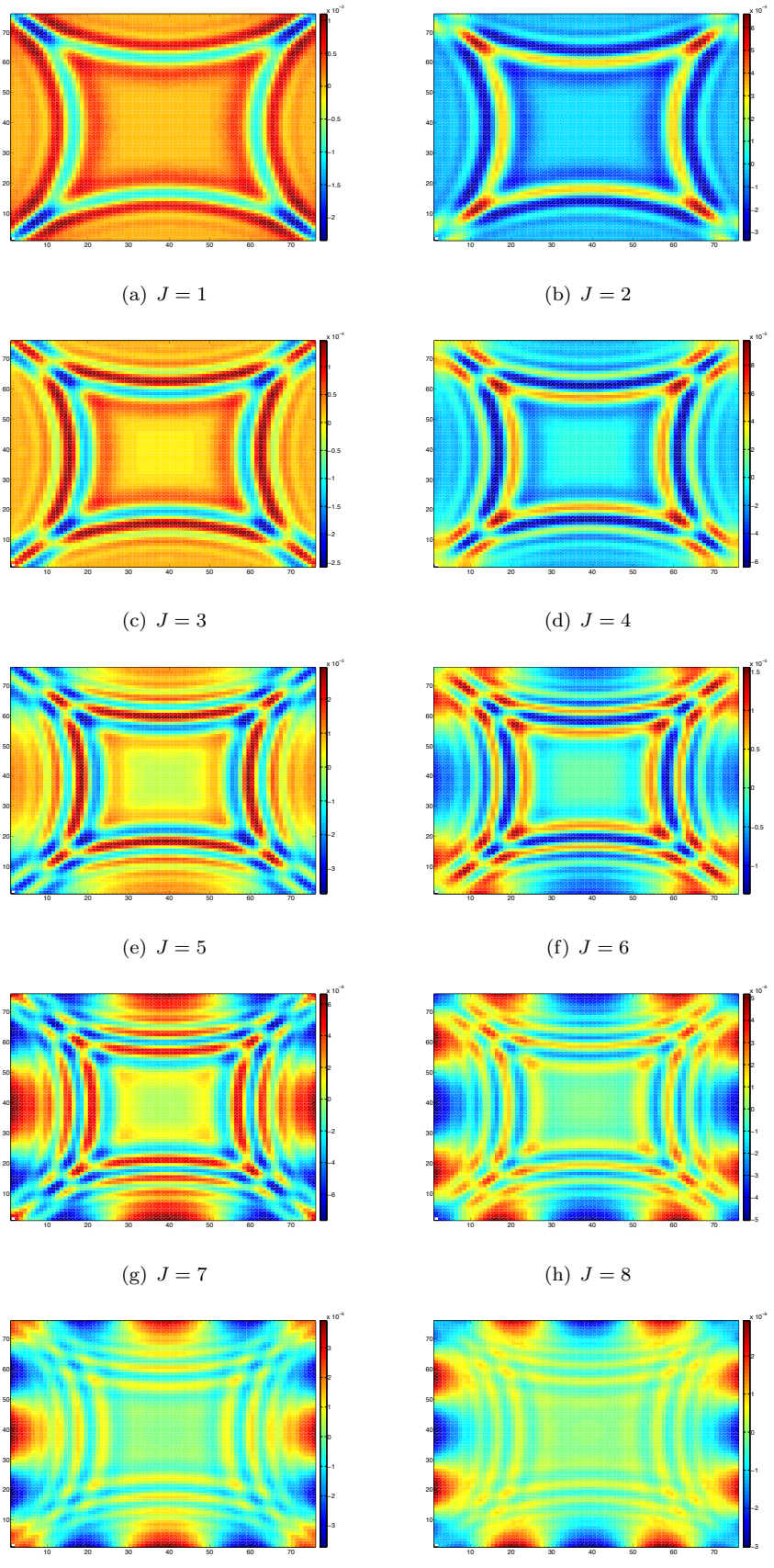
J	E_h	E_u	E_v
1	8.151605×10^{-1}	9.330289×10^{-1}	9.330289×10^{-1}
2	2.285768×10^{-1}	3.040398×10^{-1}	3.040398×10^{-1}
3	9.771136×10^{-2}	1.225036×10^{-1}	1.225036×10^{-1}
4	4.076974×10^{-2}	5.403725×10^{-2}	5.403725×10^{-2}
5	1.799678×10^{-2}	2.652240×10^{-2}	2.652240×10^{-2}
6	8.446164×10^{-3}	1.502765×10^{-2}	1.502765×10^{-2}
7	4.377410×10^{-3}	9.499752×10^{-3}	9.499750×10^{-3}
8	2.547459×10^{-3}	6.470463×10^{-3}	6.470475×10^{-3}
9	1.657524×10^{-3}	4.667585×10^{-3}	4.667606×10^{-3}
10	1.158764×10^{-3}	3.527493×10^{-3}	3.527498×10^{-3}

Table 2.9: Error norms for experiment 1, using the automatically chosen wavespeeds to generate the NRBC coefficients.

to set $J = 8$, at least for such a small grid where the boundary condition constitutes such a significant proportion of the numerical work.

Table 2.8 shows the results of experiment 2, which is run with the same parameters as in experiment 1, except that now $f = 0.5$. The results are very similar, and demonstrate the effectiveness of the NRBC even for low order.

Slightly disappointing are the results of the same experiments performed using wavespeed coefficients chosen by the automatic procedure suggested in [30]. Table 2.9 shows the field errors obtained in experiment 1, and table 2.10 shows the field errors obtained in experiment 2. An improvement as the order of the scheme is increased is still in evidence (as is proven must be the case), though not as spectacularly as with the basic choice of $c_j = 1 \ \forall j$. It is unclear why this should be the case, as the reasoning behind the procedure seems justified, although some of the choices behind it are arbitrary ones. The error between the computed height field and the reference height field are shown for all orders of J in figure 2.22. Hence we have shown that application of the Givoli & Neta procedure for determining the arbitrary wavespeeds to Dea's application implementation of the Higdon NRBC with a staggered scheme leads to less than optimal results.



(a) $J = 1$ (b) $J = 2$

(c) $J = 3$ (d) $J = 4$

(e) $J = 5$ (f) $J = 6$

(g) $J = 7$ (h) $J = 8$

(i) $J = 9$ (j) $J = 10$

Figure 2.22: Error between computed and reference height fields at the final time step, for several orders of the boundary condition, using automatically selected coefficients c_j.

J	E_h	E_u	E_v
1	5.505955×10^{-1}	4.634918×10^{-1}	4.634918×10^{-1}
2	1.584828×10^{-1}	1.490634×10^{-1}	1.490634×10^{-1}
3	6.874877×10^{-2}	6.064418×10^{-2}	6.064418×10^{-2}
4	2.888927×10^{-2}	2.724628×10^{-2}	2.724628×10^{-2}
5	1.312785×10^{-2}	1.353180×10^{-2}	1.353180×10^{-2}
6	6.386657×10^{-3}	7.693942×10^{-3}	7.693942×10^{-3}
7	3.421420×10^{-3}	4.853541×10^{-3}	4.853541×10^{-3}
8	2.025293×10^{-3}	3.294705×10^{-3}	3.294705×10^{-3}
9	1.321894×10^{-3}	2.365645×10^{-3}	2.365659×10^{-3}
10	9.238736×10^{-4}	1.779430×10^{-3}	1.779703×10^{-3}

Table 2.10: Error norms for experiment 2, using the automatically chosen wavespeeds to generate the NRBC coefficients.

2.4.9 Stability

We now carry out a stability analysis for the scheme. As pointed out by Dea [15], the staggered scheme involving a Higdon NRBC has not been subject to a stability analysis. In §2 Dea goes on to note that previous analyses of discretised systems employing a Higdon NRBC [14, 16] have failed to notice a subtle distinction between the way that the interior velocity values and the interior height values are calculated in the derivation. Such derivations assumed that the velocity fields on the boundary are computed using the interior difference scheme, whereas in fact this is not true. Hence, even when $f = 0$, we cannot simply rely on the stability result for the discretised 1-d Klein-Gordon equation with a Higdon NRBC, as this scheme is not matched perfectly by the one considered here. However, it is encouraging that no serious stability issues were observed in Dea's experiments.

Analysis of the staggered scheme is not possible analytically, and instead we pursue a semi-numerical route to determining any conditions on its stability. First, let us write the values of all three fields at a particular time level, in a single vector $\boldsymbol{\phi}$

$$\boldsymbol{\phi}^n = (\boldsymbol{\eta}^n, \mathbf{u}^{n-\frac{1}{2}}, \mathbf{v}^{n-\frac{1}{2}})^T, \tag{2.120}$$

where the vectors $\boldsymbol{\eta}^n$, $\mathbf{u}^{n-\frac{1}{2}}$ and $\mathbf{v}^{n-\frac{1}{2}}$ contain the relevant solutions evaluated at every point on the respective grids. Both their ordering within $\boldsymbol{\eta}_n$ and their internal ordering is unimportant for now. Now, the solution is evolved by using the interior scheme and the boundary conditions, to determine $\boldsymbol{\eta}^{n+1}, \mathbf{u}^{n+\frac{1}{2}}$ and $\mathbf{v}^{n+\frac{1}{2}}$ from $\boldsymbol{\eta}^{n-k}, \mathbf{u}^{n-\frac{1}{2}-k}$ and $\mathbf{v}^{n-\frac{1}{2}-k}$, where k runs from zero to J, the order of the NRBC. Hence, by constructing a storage vector $\bar{\boldsymbol{\phi}}^n = ((\boldsymbol{\phi}^n)^T, (\boldsymbol{\phi}^{n-1})^T, \ldots, (\boldsymbol{\phi}^{n-J})^T)^T$, we may write the update rule that evolves the interior solution and applies the boundary conditions, as a matrix multiplication

$$\bar{\boldsymbol{\phi}}^{n+1} = \mathbf{A}\bar{\boldsymbol{\phi}}^n, \tag{2.121}$$

where $\mathbf{A}$ — whose form is to be determined — is the *update matrix*. The storage vector $\bar{\boldsymbol{\phi}}$ may actually be trimmed a little by remembering that we do not require the solution on the entire grid for all J levels in time; we only need store those values that will contribute to the boundary condition, which cuts the storage requirement roughly in half.

This reformulation reduces the problem of assessing the stability of the discrete scheme to a question of the **behaviour or the powers of the update matrix A, specifically, whether it is power bounded** (as defined below). This method of analysing the stability of a finite-difference scheme, though well established, has been applied relatively infrequently in the past as noted by Trefethen in chapter 34 of [74]. Details of the construction of the update matrix are given in Appendix A.

2.4.10 Power bounds for the update matrix

A most severe instability will result if some norm of powers of $\mathbf{A}$ grows unchecked to infinity. For this reason, we require that the matrix at least be *power bounded*, i.e.

$$p(\mathbf{A}) = \sup_{k \geq 0} ||\mathbf{A}^k|| < \infty, \tag{2.122}$$

where $||\cdot||$ is the matrix norm induced by some vector norm. This is equivalent to requiring

$$\rho(\mathbf{A}) < 1 \tag{2.123}$$

where $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, i.e. a simple test of stability is to compute the largest eigenvalue of $\mathbf{A}$.

Note that (2.122) does not rule out the possibility of the matrix norm growing very large at some point, before settling down to its asymptotic form. Such transient, "mild" instabilities are not revealed by the spectrum of $\mathbf{A}$, but may instead be revealed via the *pseudospectrum* of $\mathbf{A}$. There are many entirely equivalent definitions of the pseudospectrum of $\mathbf{A}$, which may be thought of as an extension of the spectrum, $\sigma(\mathbf{A})$. For $\mathbf{A} \in \mathbb{C}^{N \times N}$, the spectrum may be defined via the resolvent norm

$$||(z - \mathbf{A})^{-1}|| = \infty \quad \text{for} \quad z \in \sigma(\mathbf{A}), \tag{2.124}$$

where z will be used as a shorthand for $z\mathbf{I}$ ($\mathbf{I}$ being the identity matrix), the meaning being made clear by the context. It is then natural to extend this definition in the following way [73, 74]:

Definition 1. *The ϵ-pseudospectrum $\sigma_\epsilon(\mathbf{A})$ of $\mathbf{A}$ is the set of $z \in \mathbb{C}$ satisfying*

$$||(z - \mathbf{A})^{-1}|| > 1/\epsilon, \tag{2.125}$$

where $\epsilon > 0$.

Intuitively, whereas eigenvalues are points in the complex z-plane at which the resolvent norm is infinite, the pseudospectrum can be thought of as a set of points, on the boundary of which the resolvent norm attains a specific value. Although the definition of the pseudospectrum leaves the choice of norm unspecified, when it comes to actually computing pseudospectra we will find that it is most advantageous to work with the spectral or 2-norm, i.e. $||\cdot||_2$. The pseudospectrum can be used to reveal much about the behaviour of a matrix, and in particular is necessary to derive an estimate of how large the norm of a matrix can grow, by virtue of the

following key result.

Theorem (Kreiss-Spijker)**.** *For any* $\mathbf{A} \in \mathbb{C}^{N\times N}$,

$$\mathcal{K}(\mathbf{A}) \leq p(\mathbf{A}) \leq eN\mathcal{K}(\mathbf{A}). \tag{2.126}$$

From this theorem we have bounds for how large the powers of $\mathbf{A}$ may grow, which are given in terms of the *Kreiss constant*, $\mathcal{K}(\mathbf{A})$.

Definition 2. *For a matrix* $\mathbf{A}$*, the Kreiss constant* $\mathcal{K}(\mathbf{A})$ *is*

$$\mathcal{K}(\mathbf{A}) = \sup_{|z|>1} (|z| - 1)\, ||(z - \mathbf{A})^{-1}||. \tag{2.127}$$

From the above theorem and definition, we can make the following link between the growth of the norms of matrix powers, and the pseudospectrum: $p(\mathbf{A})$ has an upper bound related to the Kreiss constant, which from definition 2 is related to the rate at which the resolvent norm increases as z comes near to an eigenvalue of $\mathbf{A}$. From this it is clear that a large value of $\mathcal{K}$ will be associated with a bulge in the pseudospectrum near a particular eigenvalue, and hence such bulges will alert us to the presence of potentially large transient growth of $||\mathbf{A}^k||$, and accordingly a quietly growing mode admitted by the difference scheme which has $\mathbf{A}$ as its update matrix.

Consider as an example, the matrix

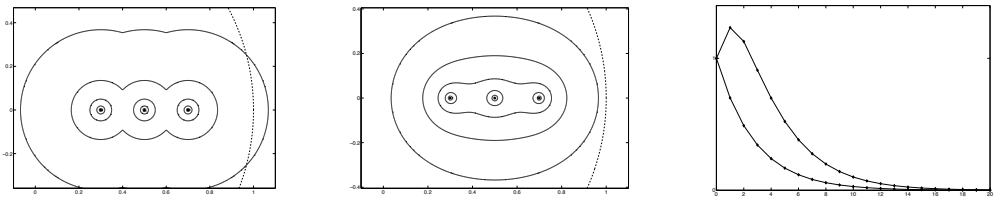
$$\mathbf{B} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{5} \\ 0 & \frac{1}{2} & 0 \\ \frac{1}{5} & 0 & \frac{1}{2} \end{bmatrix}.$$

This matrix has eigenvalues of 0.7, 0.5 and 0.3, so is power bounded. The matrix is also normal, i.e. $\mathbf{B}\mathbf{B}^\dagger = \mathbf{B}^\dagger\mathbf{B}$, where $\dagger$ denotes the Hermitian transpose. For normal matrices, the pseudospectra form neat loci around the eigenvalues, and we should not observe any bulges. We expect the norms of powers of $\mathbf{B}$ to be strictly decreasing. Figures 2.23a illustrates the pseudospectrum of $\mathbf{B}$, and figure 2.23c shows the behaviour of $||\mathbf{B}^k||$; both are as expected. Now we perturb one entry (in the first row) of $\mathbf{B}$ to form

$$\mathbf{C} = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{5} \\ 0 & \frac{1}{2} & 0 \\ \frac{1}{5} & 0 & \frac{1}{2} \end{bmatrix}.$$

Clearly this matrix has the same eigenvalues as $\mathbf{B}$, but now the matrix is nonnormal. Looking at the pseudospectrum in figure 2.23b, we observe a protrusion around the centre eigenvalue, and from figure 2.23c see that this is associated with a hump in the norms of the powers of $\mathbf{C}$.

The literature on the use of pseudospectra for determination of stability of numerical discretisations of PDEs is confined to a few papers, indicating that the method is underexploited. In particular a trio of papers by Zingg and his collaborators [91, 89, 90] give many examples of



(a) $\sigma_\epsilon(\mathbf{B})$ (b) $\sigma_\epsilon(\mathbf{C})$ (c) $||\mathbf{B}^k||$ (lower curve), $||\mathbf{C}^k||$ (upper curve).

Figure 2.23: Pseudospectra for the two example matrices (with the unit circle shown by a dashed line), together with their power norms.

systems to which the method is applied, though many of these are cases where stability results are easily obtained by well established methods. Hence, the present work is one of very few in which stability is investigated in such a way.

2.4.11 Computing pseudospectra

Definition 1 of $\sigma_\epsilon(\mathbf{A})$ in terms of the resolvent norm suggests calculating pseudospectra by computing a series of inverses and their norms for different values of z, chosen to lie on a regular grid in the complex plane. This procedure is costly for large matrices. To make progress, we make the choice $||\cdot|| = ||\cdot||_2$, then

$$||(z-\mathbf{A})^{-1}|| = [s_{\min}(z-\mathbf{A})]^{-1}; \tag{2.128}$$

here $s_{\min}(\cdot)$ is the minimum singular value. Hence, we may employ the following definition

Definition 3. *With the choice $||\cdot|| = ||\cdot||_2$, the ϵ-pseudospectrum $\sigma_\epsilon(\mathbf{A})$ of $\mathbf{A}$ is the set of $z \in \mathbb{C}$ satisfying*

$$s_{min}(z-\mathbf{A}) < \epsilon. \tag{2.129}$$

where $\epsilon > 0$.

Computing pseudospectra via singular values is still a costly procedure, and also wastes effort by computing *all* of the singular values whereas only one is required. A major improvement may obtained by using an iterative method such as inverse iteration (which may be preceded by triangularisation, as proposed by Lui [57]) or Lanczos iteration.

We implement the *boundary tracing* method proposed by Kostin [18] and subsequently refined into a robust algorithm by Brühl [9]. If we wish to trace the boundary of $\sigma_\epsilon(\mathbf{A})$ for a particular ϵ, then using a grid is not the best option, since most points will lie away from this particular curve. Brühl's algorithm for tracing a pseudospectral boundary is summarised below.

Algorithm 4 Kostin-Brühl boundary tracing method

Begin with an initial guess $\hat{z}_0$ for a point on the pseudospectral boundary to be traced, $\partial\sigma_\epsilon(\mathbf{A})$.

From this point, use a root finding (e.g. Newton) method to find (to tolerable accuracy) $z_0 \in \partial\sigma_\epsilon(\mathbf{A})$.

for all n **do**

Compute a nearby point $\hat{z}_{n+1}$ a distance of τ from z_n in the direction of *least* descent of the function $||(z-\mathbf{A})|| = \sigma_{\min}(z-\mathbf{A})$.

Take a single Newton step to correct $\hat{z}_{n+1}$ to z_{n+1}.

end for

The direction of least descent of the resolvent norm is orthogonal in the complex z-plane to the direction of steepest descent, which itself may be shown to be

$$\nabla f(x,y) = (\mathrm{Re}\,(\mathbf{v}^*\mathbf{u}), \mathrm{Im}\,(\mathbf{v}^*,\mathbf{u})), \tag{2.130}$$

where $\mathbf{u}$ and $\mathbf{v}$ are the corresponding left and right singular vectors of $\sigma_{\min}(z-\mathbf{A}) = \sigma_{\min}(x+\mathrm{i}y-\mathbf{A})$. Hence, the direction in the complex plane from z_n to $\hat{z}_{n+1}$ is $\mathrm{i}\mathbf{v}^*\mathbf{u}$. The step size τ must be chosen small enough to give an accurate picture of the boundary. It is possible that in some cases the method fails, e.g. where there are sharp corners in $\partial\sigma_\epsilon(\mathbf{A})$, but this can usually be overcome by reducing τ. The method can even be improved by allowing τ to vary, so that

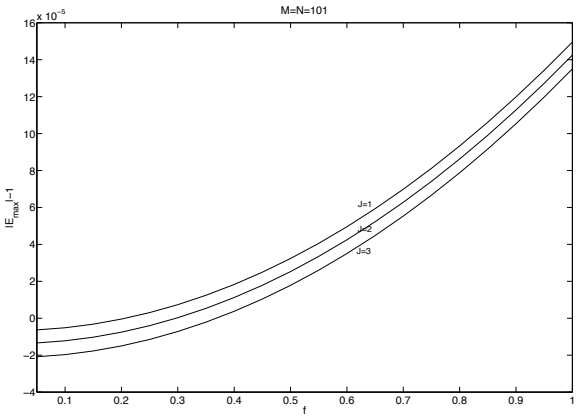


Figure 2.24: Maximum eigenvalues obtained using a fixed grid with $M = N = 101$, together with $\Delta x = \Delta y = 0.1$ and $\Delta t = 0.0177$, with $J = 1, 2, 3$.

it is reduced when the gradient $\nabla f(x, y)$ becomes large.

When plotting pseudospectral boundaries, the pseudospectrum may be first visualised roughly using grid methods for sparse matrices, as implemented in the software *Eigtool*[3] [88]. Then the boundary tracing method (which we have implemented in MATLAB) can be applied to draw a particular level curve.

2.4.12 Stability results

First, we examine the spectral problem in order to determine how stability is related to the order of the boundary condition, J, and the Coriolis parameter f. We investigate the power boundedness of the update matrix by determining its largest eigenvalues for various values of f and J. A grid step of $\Delta x = \Delta y = 0.1$ was used, together with $\Delta t = 1/4\Delta t_{\mathrm{CFL}} = 0.0177$. The scheme was tested for $J = 1, 2$ and 3, with $f \in [0, 1]$ on a grid comprising 101×101 points. The adjustable wavespeeds were all set to one for simplicity. Figure 2.24 shows that the stability threshold is crossed as f is increased, and that the higher the order of the scheme, the larger the value of f that it will tolerate. The stability threshold also varies with the grid size, as shown by figures 2.25–2.27 which plot $|\sigma_{\max}|$ against f and the grid size M for square grids for schemes with $J = 1, 2$ and 3. The results are probably best illustrated by plotting the contours of $|\sigma_{\max}| - 1$, as in figures 2.28 –2.30, which show clearly the stability threshold boundary.

By choosing a set of parameters that yields a scheme that is absolutely stable, we may search for less severe instabilities by computing the pseudospectra. Figure 2.31 shows the ϵ-pseudospectral boundaries computed (using the Kostin-Brühl method) for $\epsilon = 0.1$ with $\tau = 0.01$, for $M = N = 31, J = 1$ and $f = 0$, $f = 0.1$ and $f = 0.5$. We notice a bulge in the pseudospectrum around the unit circle in the direction of the real axis for all three values of f, indicating the existence of a mild instability, although all three plots are very similar. Similarly, figure 2.32 shows the results for the same grid with $f = 0$, with $J = 1, 2, 3$. Again we note that

[3]Eigtool is available as a MATLAB package, and implements gridded methods to compute pseudospectra directly or using iterative methods and methods for sparse matrices. These are combined with an interactive gui for convenience.

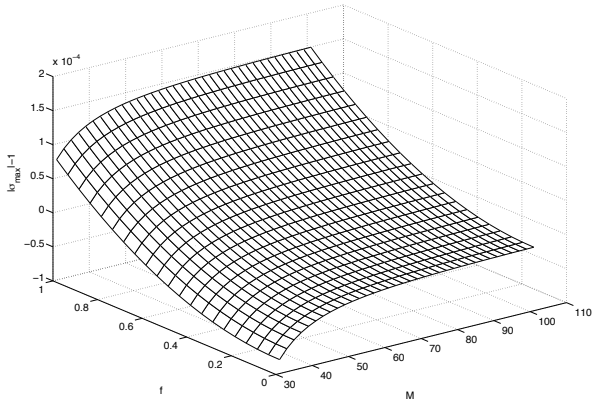


Figure 2.25: Eigenvalue surface obtained with $J = 1$.

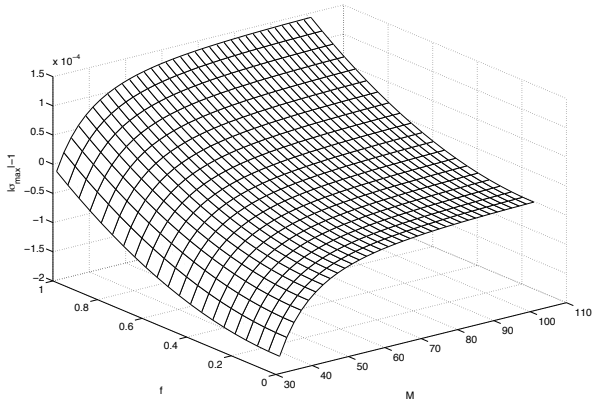


Figure 2.26: Eigenvalue surface obtained with $J = 2$.

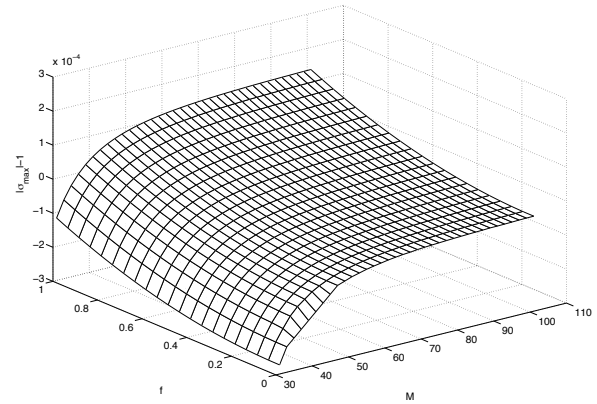


Figure 2.27: Eigenvalue surface obtained with $J = 3$.

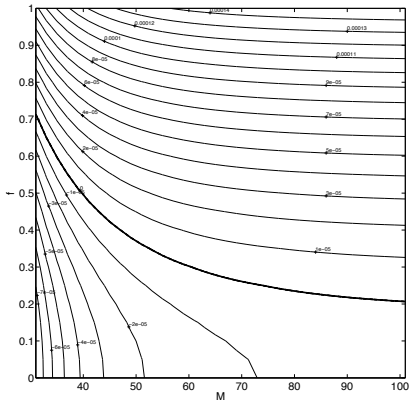


Figure 2.28: Contours of $|\sigma_{\max}| - 1$ obtained for $J = 1$. The boundary between stable and unstable regions is marked by the thick black line.

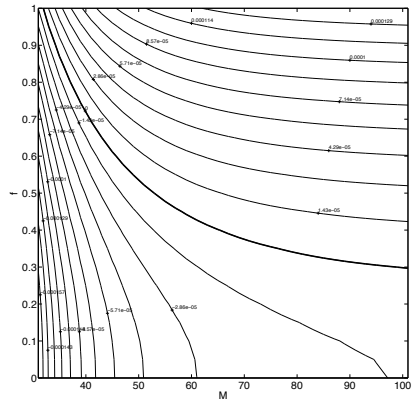


Figure 2.29: Contours of $|\sigma_{\max}| - 1$ obtained for $J = 2$; the thick line corresponds to $|\sigma_{\max}| = 1$.

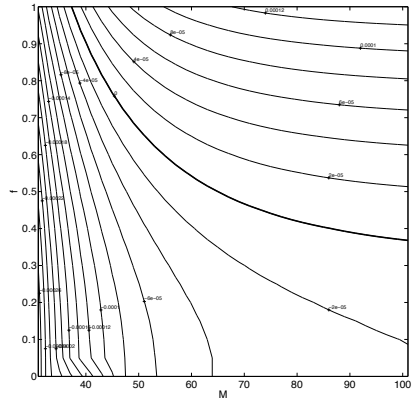
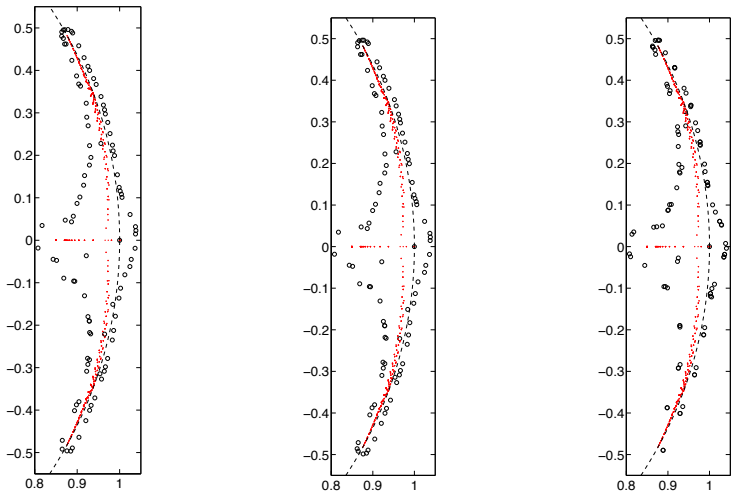
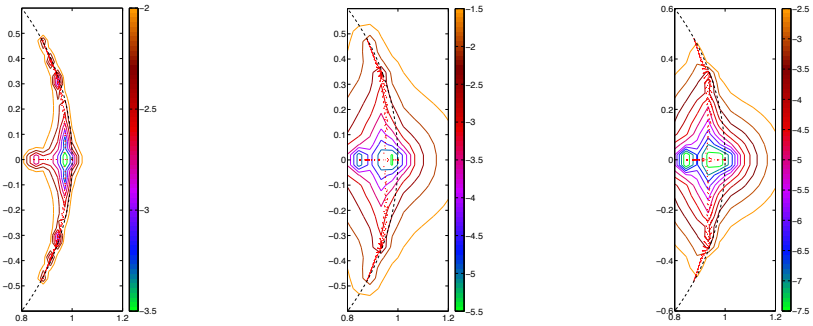


Figure 2.30: Contours of $|\sigma_{\max}| - 1$ obtained for $J = 3$; the thick line corresponds to $|\sigma_{\max}| = 1$.



(a) $f = 0$ (b) $f = 0.1$ (c) $f = 0.5$

Figure 2.31: Contours of pseudospectra (black circles) found using the boundary tracing algorithm, and eigenvalues (red dots) for $M = N = 31$ and $J = 1$ for varying Coriolis parameter. The unit circle is represented by the dotted line.



(a) $J = 1$ (b) $J = 2$ (c) $J = 3$

Figure 2.32: Pseudospectra (solid lines) found using Eigtool, and eigenvalues (red dots) for $M = N = 31$ and $f = 0$, for varying scheme order. Scales are $\log_{10}(\epsilon)$, with the unit circle represented by the dotted line.

the pseudospectrum bulge outside the unit circle.

2.5 Conclusions

In this section we have applied the Givoli-Neta algorithm for choosing free parameters in a Higdon higher-order boundary condition, to a staggered-time/staggered-space scheme, and have shown that the results are less favourable than with the simpler choice of $c_j = 1$. A novel and practical algorithm was derived for implementing the scheme. Continuing work along the lines of a suggestion by Dea, we showed using a spectral/pseudospectral technique that the scheme may be absolutely unstable for certain choices of grid size and Coriolis parameter, and that it may exhibit milder instabilities even if it is absolutely stable. This is one of only a few examples in the literature of the application of such a technique.

Chapter 3

Maintenance of balance by wave radiation in a toy model

3.1 Wave radiation and balance

The ocean is an example of a system that evolves on widely separated timescales: large scale circulation takes place over a relatively long timescale, typically of many days, whilst inertia-gravity waves are much faster, with timescales on the order of minutes. Faster still are acoustic waves, though these are almost always deliberately neglected by ocean models by introducing an anelastic or incompressible constraint in the fluid equations. Given any such system whose evolutionary timescales are so disparate, a natural question to ask is to what extent each branch of the dynamics interacts with the other; in the case of the ocean, such a question remains an open one. In geophysical flows, the timescale separation between the fast waves and slow, balanced motion is indicated by the Rossby number, U/fL.

In many such models, the fast modes of activity are very weak; typically, the dynamics means that the energy forced into the system is restricted to exciting slow motion, leaving the fast modes only weakly excited. Hence, it is natural to attempt to reduce the model's dynamics onto a *slow manifold* [44], a structure in the state space on which the dynamics is slow, but closely approximates the full dynamics. Such a restriction constitutes a *balanced model* [81], with classic examples of balanced models in geophysical fluids including the quasigeostrophic equations and the semi-geostrophic equations [22]. In the ocean, if slow motion takes place on a typical non-dimensionalised timescale of $T = \mathcal{O}(1)$, then the fast inertia-gravity waves can be shown to have a timescale of $\mathcal{O}(R_o^{-1})$, where $R_o = U/fL$ is the Rossby number, and U, f and L are a characteristic velocity, Coriolis parameter, and length scale, respectively. Motions which lie close to this slow "fuzzy" manifold (which sometimes earns it the name *quasimanifold* [25]) are said to be balanced, with any deviations corresponding to fast, unbalanced motion. Balanced geophysical models capture the slow timescale dynamical evolution, whilst excluding most of the fast gravity wave content. By an iterative procedure, slow manifolds can be derived that are successively closer to being invariant. Denoting the fast variables by u_f and the slow variables by u_s, it is possible to find a hierarchy of relations such that

$$\mathbf{u}_\mathrm{f} = \epsilon F_1(\mathbf{u}_\mathrm{s}) + \epsilon^2 F_2(\mathbf{u}_\mathrm{s}) + \epsilon^3 F_3(\mathbf{u}_\mathrm{s}) + \ldots, \tag{3.1}$$

where $\epsilon \ll 1$. A useful reference on the derivation of slaving relations and balanced models may be found in [84]

In addition to its internal dynamics, the ocean is constantly being forced by wind stress. This random forcing is generally slow and localised in space. It is now appropriate to combine all of these features of the ocean and ask two important questions using a model system. First, given that there is a forcing with a particular random structure, how does this energy manifest itself in the interior dynamics? Second, given that energy is injected into our system, what role (if any) does the fast motion play in maintaining balance? In other words, we wish to know whether there is a mechanism whereby excess energy is shed by the system in the form of gravity waves which radiate to infinity, i.e. a balanced state is preserved by dissipation. A major question in oceanography surrounds the mechanism by which the energy budget of the global circulation is closed. Specifically, there is speculation concerning the transfer of energy from mesoscale eddies down to the scale of internal waves. Spontaneous generation of inertia-gravity waves has been proposed as a "missing link" in this budget based on two layer rotating annulus experiments [86], however the model which we study suggests that, at least for the range of Rossby numbers encountered in the ocean, such an effect is too small to account entirely for the transfer of energy. Similar questions relating to the atmosphere remain open [59]. Typically in the ocean, the timescale of the IGW radiation and the forcing are widely separated, with the effect that the fast modes are only weakly excited, making the ocean an ideal candidate for representation by a balanced model. We answer these questions using a novel forced dissipative system, where dissipation is through wave radiation.

The variety of timescales that are present makes for an interesting system. The intention is to explore different types of forcing and examine the resulting dynamics. The forcing is a smooth random function, whose dominant timescale and amplitude may be varied. The expectation given the wide timescale separation of the slow and fast modes is that using slow forcing, the fast dynamics is dominated by the high frequency components of the slow dynamics, and is very little influenced by the external forcing. Conversely, when the forcing is fast, the forcing and fast modes should show a clear degree of synchronicity. The effect of varying amplitude may be studied, in an effort to extract a scaling argument for the flux of the emitted waves.

Of course, accurate modelling of the ocean is a very involved procedure, so instead a toy mixed ODE/PDE model is used, for which balance relations may be constructed in a manner that is straightforward. Such toy models provide the basis for much of the work in the remainder of this thesis, and many of the techniques from chapters 2 and 3 may be brought to bear in chapter 4. Although not physically realistic, the model broadly reproduces many aspects of the ocean dynamics that we seek to reproduce. For example, large timescale separation, a dispersion relation identical to that of the shallow water equations, and the ability to be forced and emit gravity waves. It should also reflect the weakness of the small fast waves, which in at least some simple geophysical models have a flux that is $\mathcal{O}(\exp(-\alpha/R_o))$ [82, 80]. Dissipation is often introduced via a damping term, which affects both the slow and fast dynamics. In contrast, the model we employ features dissipation by wave radiation, which affects almost exclusively the fast inertia-gravity-type waves.

The results of this section are as follows. First, in §3.1.1 a model due to work by Lorenz, Krishnamurthy and Vanneste is introduced. This model is a coupled ODE/PDE system, and has some qualitative features which are pertinent to the wind-forced ocean problem. Next, the model is extended by the addition of a smooth random forcing term to the equations for the slow variables. Following on from work in [79], a high order balanced solution is sought using an asymptotic expansion in §3.1.3, in order to separate the fast motions from the underlying slow motions. The model has the potential to answer many questions surrounding various scaling relations between different parameters and quantities of interest, and such questions are explored. In §3.2 the balance relations are employed in a numerical simulation of the system (incorporating an exact, discretely non-reflecting boundary condition, as derived in §2.3.3 of chapter 2) in order to examine how wave radiation is affected by various types of forcing.

3.1.1 A family of toy models

In a well-known paper [55], Lorenz considered a simple five component ODE model. This was derived from the shallow water equations on the f-plane, by expanding each field as a double Fourier series, and truncating to leave only three modes (as carried out in [54]). Next, all but one nonlinear terms are discarded, and the quasigeostrophic approximation is made for two of the three Fourier modes. Subsequently, Lorenz and Krishnamurthy [56] reformulated this model to give the system that now bears their name:

$$\dot{u} = -vw + bvy, \tag{3.2}$$

$$\dot{v} = wu - buy, \tag{3.3}$$

$$\dot{w} = -uv, \tag{3.4}$$

$$\dot{\delta}x = -y, \tag{3.5}$$

$$\dot{\delta}y = x + b\delta uv; \tag{3.6}$$

here $\delta = \epsilon b/\sqrt{1+b^2} \ll 1$ is a small parameter, whose presence sets apart x and y from the other variables by making them fast. Up to this point, the analogy with an ocean model only extends as far as there being a time scale separation between the slow variables u, v and w, and the fast variables x and y which can be seen to describe a linear oscillator coupled to the slow variables. What is lacking is some mechanism to support dispersive waves, which are ideally allowed to radiate their energy to infinity. Although the equations are far removed from any realistic fluid model, b can be seen to correspond to a rotational Froude number

$$b = F_r = fL/\sqrt{gH}, \tag{3.7}$$

and it is useful to define a Rossby number $\epsilon = R_o = U/fL$. Vanneste studies two regimes: small Froude number $b \ll 1$, corresponding to Lighthill radiation, and small Rossby number $\epsilon \ll 1$; we consider the latter, being most relevant to the ocean.

In a more realistic fluid setting such as the rotating shallow water equations, there are two regimes of interest. To see their differences, consider the dispersion relation for the SWEs:

$$\omega^2 = \epsilon^{-2}(k^2/b^2 + 1); \tag{3.8}$$

In the first regime with $R_o = \epsilon \ll 1$ and $F_r = b = \mathcal{O}(1)$, the presence of the $+1$ term inside the bracket ensures waves are dispersive. This is even more apparent for small wavenumbers, where the frequency is bounded from below. Hence, there is a time scale separation between the slow, balanced motion on the slow manifold, and even the lowest frequency inertia-gravity waves. In the event, the frequency gap means that wave radiation is exponentially weak in the Rossby number [80]. However, consider reversing the situation, and taking $\epsilon = \mathcal{O}(1)$ and $b \ll 1$; now the frequency gap between the slow balanced motions and the long gravity waves is shrunk, and Lighthill-type radiation occurs.

Vanneste [79] made a further modification to the Lorenz-Krishnamurthy model. Instead of the fast variables being uni-variate, they are made functions of both t and a spatial coordinate s. As such, they describe the dynamics of a dispersive string coupled to the slow variables via a localised function $f(s)$. The system — which will be referred to as the extended Lorenz-Krishnamurthy (eLK) system — may be expressed as follows:

$$\dot{u} = -vw + v\int f(s)y(s,t)ds, \tag{3.9}$$

$$\dot{v} = wu - u\int f(s)y(s,t)ds, \tag{3.10}$$

$$\dot{w} = -uv, \tag{3.11}$$

$$\epsilon x_t = -y, \tag{3.12}$$

$$\epsilon y_t = x - x_{ss}/b^2 + \epsilon f(s)uv. \tag{3.13}$$

By taking the time derivative of the final two equations, neglecting coupling and rearranging, the fast variables can be shown to satisfy the familiar 1-d Klein-Gordon equation (for which we have the numerical tools developed earlier at our disposal). Hence the dispersion relation for waves on the string is

$$\omega^2 = \epsilon^{-2}(1 + k^2/b^2) = f^2L^2/U^2 + gH/U^2k^2, \tag{3.14}$$

which matches that of the shallow water equations. In the model, $s \in (-\infty, \infty)$, and the function $f(s)$ describes the spatially varying coupling between this infinite string and the slow variables. In [79], the chosen function is of the form

$$f(s) = \frac{a}{(2\pi)^{1/2}}\frac{d}{ds}e^{-s^2/2} = -\frac{as\, e^{s^2/2}}{(2\pi)^{1/2}}, \tag{3.15}$$

where a is a constant. One appealing feature of such a function is that it is odd, and hence the string problem may be restricted to the half line $s \in \mathbb{R}^+$. As was pointed out in [79], an analysis of the model for $b \ll 1$ is a cartoon of that developed by Ford, McIntyre and Norton in [25], for the study of balance in the rotating SWEs. In this framework, the coupling function $f(s)$ may be thought of as an analogue of the PV distribution in shallow water.

In this work, the eLK model is extended once more by the addition of a random forcing term to the equations for the slow variables. Such an extension was suggested in [79] as a basis for future work, though not pursued. Before introducing the forcing, the equations (3.9)–(3.13) can be manipulated to reveal their Hamiltonian nature, as explained fully in [79]. Introducing ϕ with

$$u = \sqrt{2}\cos\phi, \quad v = \sqrt{2}\sin\phi, \tag{3.16}$$

and making the change of variable

$$\theta = \phi - \epsilon \int f(s)x(s,t)ds \tag{3.17}$$

yields the following system of four equations:

$$\dot{\theta} = w, \tag{3.18}$$

$$\dot{w} = -\sin\left[2\theta + 2\epsilon \int f(s)x(s,t)ds\right], \tag{3.19}$$

$$\epsilon x_t = -y, \tag{3.20}$$

$$\epsilon y_t = x - x_{ss}/b^2 + \epsilon f(s)\sin\left[2\theta + 2\epsilon \int f(s)x(s,t)ds\right]. \tag{3.21}$$

In equations (3.18) and (3.19) for the slow variables (which will be forced), we recognise the description of a nonlinear pendulum. The form of the forcing is discussed more fully in §3.1.2 on numerical methods.

The model is a novel one in that it allows the study of forced, dissipative regimes, where the *dissipation is mostly restricted to the fast waves.* Perhaps the key question is whether this model exhibits the striking feature of spontaneously generated inertia-gravity waves — waves that are exponentially small in Rossby number, with a wave flux F that behaves like $\exp(-\alpha/\epsilon)$, where α is some constant. The numerical model may be used to test such a hypothesis; even better would be the ability to predict the constant α governing the form of the flux, using knowledge of the balanced part of the solution only. To this end, consider as an example the simple forced harmonic oscillator:

$$\epsilon^2 \ddot{x}(t) + x(t) = g(t), \tag{3.22}$$

where $g(t)$ is some forcing function that satisfies $g \to 0$ as $t \to \pm\infty$. In this case, the solution for the entire system is trivial to write down using Green's functions, under the additional assumption that both $x \to 0$ and $\dot{x} \to 0$ as $t \to -\infty$:

$$x = \frac{1}{\epsilon}\int_{-\infty}^{t} g(t')\sin\frac{t-t'}{\epsilon}dt'. \tag{3.23}$$

The solution may exhibit oscillations if $g(t),\ t \in \mathbb{C}$ possesses poles. The strength of such exponentially small oscillations is governed by the proximity α of the poles to the real axis. For this system at least, it is clearly a simple matter to predict the exponential dependence of the flux, by first finding poles of $g(t)$. Indeed, it is possible to show as in [80] that if the forcing function has poles located at $t = \alpha + \mathrm{i}\beta$ and $\bar{t}$, then a contour integral of the form illustrated in figure 3.1 gives an estimate for the fast oscillations of the form

$$x_{\mathrm{f}} \sim -2\pi\epsilon^{-1}\mathrm{e}^{-\beta/\epsilon}\mathrm{Im}\left\{\mathrm{i}a\mathrm{e}^{\mathrm{i}(t-\alpha)/\epsilon}\right\}. \tag{3.24}$$

For the forced eLK model, there is no closed form solution for $x(s,t)$, so there is no convenient way of inferring the parameter α from the equations. A success of the model would be the ability to link the balanced part of the solution to the flux **without knowing what form the forcing takes**. Our idea is to examine the *power spectrum* of the pendulum motion, which is controlled by the forcing and, then use this information directly to say something about the

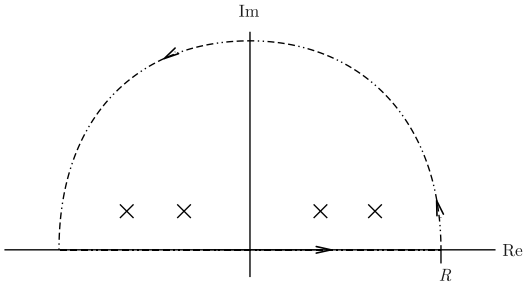


Figure 3.1: The notional integration contour for the evaluation of the Fourier transform in (3.25), with poles represented by crosses.

flux by estimating the parameter α.

To see how α may be related to the balanced motion via the forcing function, consider the power spectrum of the forcing function $g(t)$:

$$\mathcal{P}_g(\omega) = |\mathcal{F}\{g(t)\}| = \left| \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega t} g(t) \right|. \tag{3.25}$$

The contour C used to evaluate the Fourier transform may be closed by the usual introduction of a semicircle of radius R, as shown in figure 3.1. To compute the integral using the residues of the enclosed poles, we require that the integrand $I(t)$ satisfies the following requirements: $I(t)$ is analytic in the upper half of the complex t plane, except for poles shown by crosses, of which none lie on the real axis; second, $R \max_{t\in C} |I(t)| \to 0$ as $R \to \infty$. Also the integrals of $I(t)$ along the half lines $(-\infty, 0]$ and $[0, \infty)$ must exist. Assuming that these conditions are satisfied, it may be shown [1] that

$$\mathcal{P}_g(\omega) \propto \exp(-\alpha\omega), \quad \omega \gg 1. \tag{3.26}$$

Hence, a viable strategy for determining the dependence of the flux on ϵ is to estimate α directly from numerical simulations using (3.26) to examine the power spectrum of the pendulum variable, $\theta(t)$.

3.1.2 Numerical techniques

The equations (3.9–3.13) are solved on a regular grid using finite differences. Time stepping is done using a second-order Adams-Bashforth scheme, and the spatial differencing is centred and second order. The discretely non-reflecting boundary condition developed in chapter 2 is added at the right-hand boundary, using a truncation length of $K = 3000$. An additive forcing term is added to the nonlinear pendulum equation, together with a Rayleigh damping that is controlled by a constant δ. A non-reflecting boundary condition is used to allow wave radiation to infinity. We employ the exact DNRBC derived in §2.3.3, computed to high accuracy using the DFT method. The time integrations are short enough to make it practically possible to use the DNRBC *without truncation*, i.e. by storing the entire history of the solution at the endpoint.

The additive forcing is provided by a smooth, Gaussian random function $g(t)$, with a specific amplitude A and correlation function[1] $C(\tau)$, defined as

$$C(\tau) = \mathbb{E}\{f(t)f(t-\tau)\}. \tag{3.27}$$

A convenient way of constructing a function with a specific correlation function is via the Fourier transform. By defining

$$\hat{C}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} C(t)\mathrm{e}^{-\mathrm{i}\omega t}dt. \tag{3.28}$$

it follows that the sought random function $g(t)$ is given by

$$g(t) = \mathrm{Re}\left\{\int \hat{C}^{1/2}\mathrm{e}^{\mathrm{i}\omega t}(dW_\omega^1 + \mathrm{i}dW_\omega^2)\right\}, \tag{3.29}$$

where dW^1 and dW^2 represent independent Brownian motions. This result may be established by direct calculation:

$$\begin{aligned}
C_f(\tau) &= \mathbb{E}\{f(t)f(t-\tau)\} && (3.30)\\
&= 1/4\mathbb{E}\left\{\left[\int \hat{C}^{1/2}(\omega)\mathrm{e}^{\mathrm{i}\omega t}(dW_\omega^1 + \mathrm{i}dW_\omega^2) + \int \hat{C}^{1/2}(\omega)\mathrm{e}^{-\mathrm{i}\omega t}(dW_\omega^1 - \mathrm{i}dW_\omega^2)\right]\right. && (3.31)\\
&\times \left.\left[\int \hat{C}^{1/2}(\omega')\mathrm{e}^{\mathrm{i}\omega'(t-\tau)}(dW_{\omega'}^1 + \mathrm{i}dW_{\omega'}^2) + \int \hat{C}^{1/2}(\omega)\mathrm{e}^{-\mathrm{i}\omega'(t-\tau)}(dW_{\omega'}^1 - \mathrm{i}dW_{\omega'}^2)\right]\right\}\\
&= \int \hat{C}(\omega)\mathrm{e}^{\mathrm{i}\omega\tau}d\omega, && (3.32)
\end{aligned}$$

having used the identity

$$\mathbb{E}\{dW^i(\omega)dW^j(\omega)\} = \delta_{i,j}\delta(\omega-\omega')d\omega d\omega'. \tag{3.33}$$

We choose a forcing function with correlation function

$$C(\tau) = b_\mathrm{f}\mathrm{sech}(a_\mathrm{f}\tau), \tag{3.34}$$

and correspondingly,

$$\hat{C}(\omega) = \mathcal{F}\{C(\tau)\} = \frac{b_\mathrm{f}}{2a_\mathrm{f}}\mathrm{sech}(\pi\omega/2a_\mathrm{f}). \tag{3.35}$$

The two parameters a_f and b_f (where f stands for forcing), control completely the correlation function.

Here the parameter a_f is the important one, as it may be thought of as a typical inverse correlation time. The formula (3.29) is trivial to implement numerically using the inverse fast Fourier transform. Remembering that it is the poles of $g(t)$ that will control the power spectrum of the pendulum, it is worthwhile considering their location. The function $\mathrm{sech}(a_\mathrm{f}t)$ has poles lying on the imaginary axis only, located at $z = 1/a_\mathrm{f}(\mathrm{i}\pi/2 + \mathrm{i}n\pi)$, $n \in \mathbb{Z}$; the dominant pole is that closest to the axis, i.e. with $n = 0$.

There are two principal diagnostics of interest: the wave flux, and the total energy of the balanced pendulum motion. For the flux, the energy in the string is easily obtained from the

[1]The correlation function is equivalently referred to as the autocorrelation.

Klein-Gordon equation by the usual mechanical process of multiplying by x_t and integrating over the domain:

$$\begin{aligned} \epsilon^2 x_{tt} + x - x_{ss}/b^2 &= 0, && (3.36)\\ \Rightarrow \int_\Omega \epsilon^2 x_{tt}x_t + x x_t ds &= [x_s x_t/b^2], && (3.37)\\ \Rightarrow \frac{dE}{dt} = \frac{1}{2}\partial_t \int_\Omega \epsilon^2 x_t^2 + x_s^2/b^2 + x^2 ds &= [x_s x_t/b^2]. && (3.38) \end{aligned}$$

The flow of energy in the string satisfies a conservation law of the form

$$\frac{\partial E}{\partial t} + \frac{\partial F}{\partial s} = 0, \tag{3.39}$$

and from this and equation (3.38) it is clear that $F = -x_s x_t/b^2$. This is approximated using centred differences; using a grid of M points in s, where $x(s,t) = x(m\Delta s, n\Delta t)$, we find

$$F = \frac{1}{\Delta t}(x^n_{M-1} - x^{n-1}_{M-1})\frac{1}{2b^2\Delta s}(x^n_M - x^n_{M-2}). \tag{3.40}$$

The pendulum energy is approximated using

$$E_{\mathrm{p}} = \left(\frac{1}{2\Delta t}(\theta^n - \theta^{n-1})\right)^2 - \cos(\phi^n) + 1, \tag{3.41}$$

where the $+1$ term is used to normalise the pendulum rest energy to zero.

In order to understand the frequency content of the balanced motion, we must estimate the power spectrum of the balanced motion:

$$\mathcal{P}_\theta(\omega) = \hat{\theta}(\omega)\hat{\theta}(\omega)^* = \frac{1}{\sqrt{2\pi}}\hat{C}_\theta(\omega). \tag{3.42}$$

Here a hat denotes a Fourier transform in time, and $C_\theta(\tau) = \theta(t) * \theta(t-\tau)$. The latter part of the result is the well known Wiener-Kinchin theorem. When the signal undergoing analysis become discrete, we are suddenly confronted with a plethora of methods for estimating a discrete form of $\mathcal{P}_\theta(\omega)$. The simplest is the *periodogram* originally applied to weather data by Schuster [70], and is the most obvious and straightforward implementation of (3.42). For a signal with N samples $x[n]$, the periodogram is defined as

$$P(\omega_k) = \left|\sum_{n=0}^{N-1} x_n \mathrm{e}^{-\frac{2\pi i n k}{N}}\right|^2. \tag{3.43}$$

Since the Wiener-Kinchin theorem also holds for the discrete case, $P(\omega_k)$ may be rewritten as the discrete Fourier transform of the discrete approximation to the autocorrelation.

The periodogram suffers from two notable defects: the first being that even as the number of samples is increased, the variance about a particular frequency is not correspondingly reduced. Second is the fact that we deal with a finite amount of data; such a finite times series may be though of as an infinite time series multiplied by a rectangular or "top hat" windowing function, which although may be very long, is nevertheless artificial. Such a multiplication in the time

domain amounts to a convolution in the frequency domain, involving the Fourier transforms of both the infinite time series, and the rectangular function whose Fourier transform is a sinc function. Hence, there will be artefacts in the power spectrum in the form of spurious side lobes surrounding genuine peaks, with their acuteness related to the shortness of the sampling window.

To counteract both issues, we employ the *Welch periodogram*[85]. In this method, a spectral estimate is computed by averaging the periodograms of multiple overlapping portions or "frames" of the original signal, counteracting the variance problem. To alleviate the windowing problem, each of the signal frames is first multiplied by a windowing function, which is carefully chosen to reduce the spectral side lobes that would otherwise be noticeable with a plain rectangular window. Let K be the total number of frames each of M samples, and R be the window "hop" size, i.e. the number of samples between the start of subsequent overlapping frames. The Welch estimated power spectrum $\mathcal{P}_{\mathrm{W}}$ can then be written as follows:

$$\mathcal{P}_{\mathrm{W}}(\omega_k) = \frac{1}{K}\sum_{m=0}^{K-1} P_{x_m,M}(\omega_k), \tag{3.44}$$

where

$$P_{x_m,M}(\omega_k) = \frac{1}{M}\left|\sum_{n=0}^{N-1} x_m[n]\mathrm{e}^{-\frac{2\pi \mathrm{i} n k}{N}}\right|^2. \tag{3.45}$$

Here $x_m[n]$ is the nth, windowed sample of the Mth frame:

$$x_m[n] = w[n]x[n+mR], \quad n = 0,1,\ldots,M-1, \quad m = 0,1,\ldots,K-1. \tag{3.46}$$

It is clear that $K = \lfloor N_{(\mathrm{total}} - M)/R\rfloor + 1$. Common choices for R include $R = \lfloor M/3\rfloor$ and $R = \lfloor M/2\rfloor$. Two common windows are the Hamming and Hann[2] windows, defined as

$$w_{\mathrm{Hamming}}[n] = 0.54 - 0.46 * \cos(2 * \pi/(n-1)n), \tag{3.47}$$

$$w_{\mathrm{Hann}}[n] = 0.5 * (1 - \cos(2 * \pi/(n-1)n)). \tag{3.48}$$

Whilst not being optimal in terms of reducing side lobes (a Slepian window offers better performance in this regard), both are reasonable choices; we use the Hamming window throughout.

Having derived a measure of the frequency content of the forcing function in terms of power, a further interesting comparison to make is that between the wave flux over time, and the energy that is injected at the pendulum frequency. It is expected that the two are correlated, and so we compare these quantities in the hope of finding $\int F(t)dt \propto \mathcal{P}_\theta(\omega = 1/\epsilon)$.

To conclude this discussion of numerical techniques, let us perform a test of some of the reasoning and procedures developed so far. Recall that we expect $\mathcal{P}_\omega(g(t))$ — the power spectrum of the random forcing function — to decay like $\exp^{-\alpha\omega}$, and additionally $\alpha \propto \mathrm{Im}\,\omega^*$, where ω^* is a pole of $g(t)$. We may test the periodogram's effectiveness at extracting a scaling law for the power spectrum, and then examine this scaling as the forcing correlation time parameter a_{f} is varied. For this test we take the number of samples $N = 200,002, \Delta t = 0.0502, b_{\mathrm{f}} = 1,$

[2]sometimes referred to erroneously as the "Hanning" window.

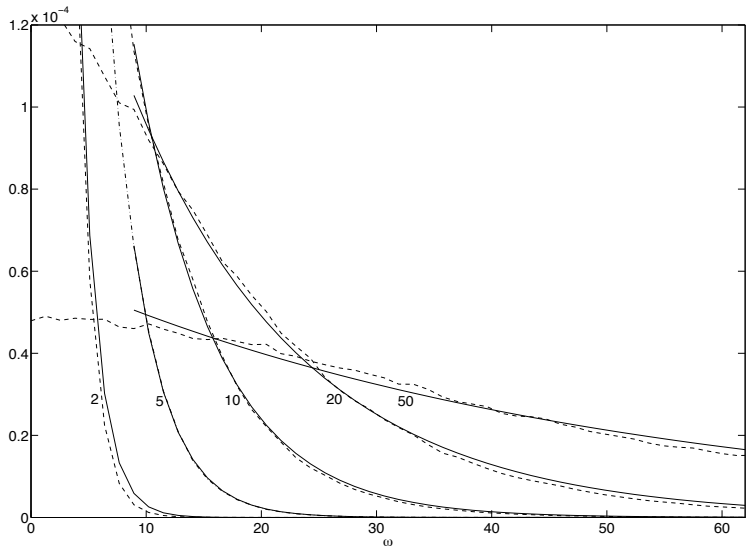


Figure 3.2: Plot of power spectra obtained by Welch periodogram (dashed lines) together with fitted exponential decay curves (solid lines), for various values of a_f (marked)

.

a_f	α
2	0.6425
5	0.3020
10	0.1448
20	0.0669
50	0.0211

Table 3.1: Results of curve fitting exercise illustrated in figure 3.2.

and employ a 101 sample Welch periodogram using a Hamming window. Figure 3.2 shows several such power spectra, each for a different value of a_f, together with a fitted decay curve (obtained by an automated procedure). Table 3.1 shows the resulting decay rates alongside a_f. The exponential fit is good, and the $a_\mathrm{f} \propto 1/\alpha$ dependence is clear.

3.1.3 Higher-order Balance relations

It is useful to be able to separate the "gravity waves" of the system (the fast motion) from the balanced part of the solution, particularly for calculating the fast wave flux close to the pendulum (whereas in the far field the motion will be almost exclusively due to the fast waves). A balanced solution of (3.9–3.13) for small Rossby number may be derived via exponential asymptotics. Specifically, expressions for slow manifolds $x_\mathrm{bal} = x(s,u,v,w)$ and $y_\mathrm{bal} = y(s,u,v,w)$ are derived, which can be successively extended to higher order. This was carried out up to zeroth order in [79], but here it is continued further. Of course, being an asymptotic expansion, there is a limiting order beyond which the solution begins to diverge, but we will not be pursuing the derivation anywhere near that far.

First, the fast balanced solutions are written as an expansion in powers of the small param-

eter, ϵ:

$$x_{\rm bal}(s,\theta,w) = \sum_{n=0}^{\infty} \epsilon^{2n+1} x_{\rm bal}^{(n)}(s,u,v,w), \tag{3.49}$$

$$y_{\rm bal}(s,\theta,w) = \sum_{n=0}^{\infty} \epsilon^{2n+2} y_{\rm bal}^{(n)}(s,u,v,w). \tag{3.50}$$

Now, equations (3.9)–(3.13) are Fourier transformed in s to yield

$$\hat{u} = -\hat{v}\hat{w} + \hat{v}\hat{f}\hat{y}, \tag{3.51}$$

$$\hat{v} = \hat{w}\hat{u} - \hat{u}\hat{f}\hat{y}, \tag{3.52}$$

$$\hat{w} = -\hat{u}\hat{v}, \tag{3.53}$$

$$\epsilon\hat{x}_t = -\hat{y}, \tag{3.54}$$

$$\epsilon\hat{y}_t = \omega^2\hat{x} + \epsilon\hat{f}\hat{u}\hat{v}. \tag{3.55}$$

Here $\mathcal{F}\{x(s,t)\} = \hat{x}(\sigma,t)$, and we have used the dispersion relation

$$\omega^2 = 1 + k^2/b^2. \tag{3.56}$$

Now, dropping all hats for convenience, let

$$x = X(u,v,w), \quad y = Y(u,v,w) \tag{3.57}$$

where upper case letters now refer to the balanced solution. From (3.54) together with (3.51)–(3.53) we find that

$$\epsilon\left(\frac{\partial X}{\partial y}(-vw+vfy) + \frac{\partial X}{\partial v}(wu-ufy) + \frac{\partial X}{\partial w}(-uv)\right) = Y, \tag{3.58}$$

and similarly from (3.55) together with (3.51)–(3.53) we find that

$$\epsilon\left(\frac{\partial Y}{\partial y}(-vw+vfy) + \frac{\partial Y}{\partial v}(wu-ufy) + \frac{\partial Y}{\partial w}(-uv)\right) = \omega^2 X + \epsilon fuv. \tag{3.59}$$

Now, substituting the expansions (3.49) and (3.50) in (3.59) and comparing powers of ϵ, we obtain

$$X_0 = -\frac{fuv}{\omega^2}, \tag{3.60}$$

and substituting this result into (3.58) yields

$$Y_0 = -\frac{fw}{\omega^2}(v^2-u^2). \tag{3.61}$$

Now we continue to higher order. Substitute (3.61) in (3.59) to give

$$X_1 = \frac{f}{\omega^4}\left[-4uvw^2 - u^3v + v^3u\right], \tag{3.62}$$

and using this result in (3.58) yields

$$Y_1 = \frac{1}{\omega^4}\left[-f^3v^4w + 2f^3u^2w - 4fv^2w^3 - 14fu^2v^2w + fv^4w - f^3u^4w + 4fu^2w^3 + fu^4w\right]. \tag{3.63}$$

We continue with this bootstrapping and find solutions up to order 3 (given in Appendix D).

Recalling that these expressions were derived by first transforming in s, it is necessary to invert the Fourier transforms. Fortunately, the results are available by analytical, rather than numerical methods. We will consider only the solution X_n. From the definitions of ϕ and θ we find

$$\begin{aligned} X_0 = -\frac{\hat{f}\hat{u}\hat{v}}{\omega^2} &= -\frac{\hat{f}\sqrt{2}\cos(\hat{\phi})\sqrt{2}\sin\hat{\phi}}{\omega^2} && (3.64)\\ &= -\frac{\hat{f}(k)\sin(2\hat{\theta})}{\omega^2}, && (3.65)\end{aligned}$$

and

$$X_1 = \frac{\hat{f}}{\omega^4}\left[-4\sin(2\hat{\phi} + 2\sin(2\hat{\phi})\cos(2\hat{\phi})\right]. \tag{3.66}$$

Recalling the form of f and inverting X_0 gives

$$\begin{aligned} x_0(s) &= 2\pi\frac{ab^2}{8}\pi\left[-2\sinh(bs)\mathrm{e}^{b^2/2} + \mathrm{erf}\left(\frac{b+s}{\sqrt{2}}\right)\mathrm{e}^{b^2/2+bs}\right. && (3.67)\\ &- \left.\mathrm{erf}\left(\frac{b-2}{\sqrt{2}}\right)\mathrm{e}^{b^2/2-bs}\right]\sin(2\theta) && (3.68)\\ &= -\frac{ab^2\mathrm{e}^{b^2/2}}{8}\left[\mathrm{erfc}\left(\frac{b+s}{\sqrt{2}}\right)\mathrm{e}^{bs} - \mathrm{erfc}\left(\frac{b-s}{\sqrt{2}}\right)\mathrm{e}^{-bs}\right]\sin(2\theta), && (3.69)\end{aligned}$$

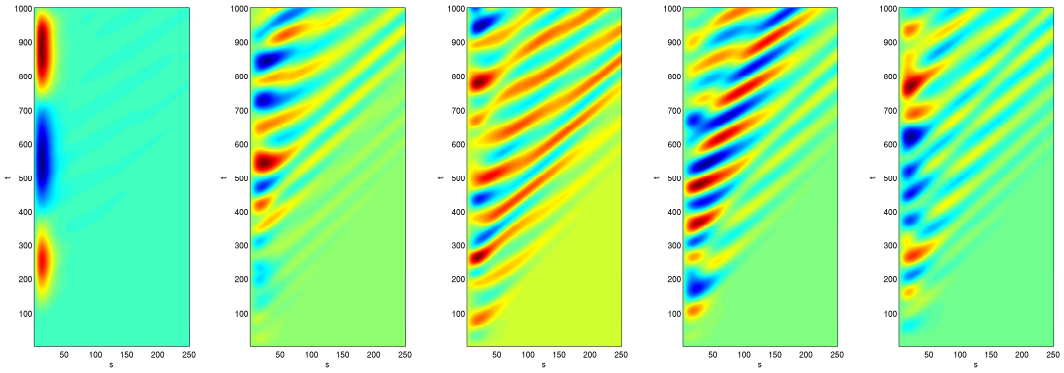
where erfc is the complementary error function. We continue with inversion of X_1:

$$\begin{aligned} \mathcal{F}^{-1}\left\{\frac{\hat{f}}{\omega^4}\right\} &= -\frac{ab^3}{4}(b\sinh(bs) + s\cosh(bs))\mathrm{e}^{b^2/2} && (3.70)\\ &+ \frac{ab^3}{8}\left[-\mathrm{erf}\left(\frac{b-s}{\sqrt{2}}\right)\mathrm{e}^{b^2/2-bs}(b-s)\right] + \frac{ab^3}{8}\left[\mathrm{erf}\left(\frac{b-s}{\sqrt{2}}\right)\mathrm{e}^{b^2/2+bs}(b-s)\right]\\ &= -\frac{ab^3\mathrm{e}^{b^2/2}}{8}\left[\mathrm{erfc}\left(\frac{b+s}{\sqrt{2}}\right)\mathrm{e}^{bs}(b+s) - \mathrm{erfc}\left(\frac{b-s}{\sqrt{2}}\right)\mathrm{e}^{-bs}(b-s)\right], \end{aligned}$$

and hence

$$\begin{aligned} x_1 = \left[-\frac{ab^3\mathrm{e}^{b^2/2}}{8}\left[\mathrm{erfc}\left(\frac{b+s}{\sqrt{2}}\right)\mathrm{e}^{bs}(b+s) - \mathrm{erfc}\left(\frac{b-s}{\sqrt{2}}\right)\mathrm{e}^{-bs}(b-s)\right]\right] && (3.71)\\ \times\left(-4\sin(2\theta)w^2 + 2\sin(2\theta)\cos(2\theta)\right) && (3.72)\end{aligned}$$

Unfortunately, it is not possible to derive a simple recurrence relation for x_n, so we continue to invert directly by way of computer algebra, up to $n = 3$; in the interests of space the resulting expressions are deferred to Appendix D. As pointed out in [79], there is nothing preventing



(a) Full, unbalanced solution. (b) Order 0. (c) Order 1. (d) Order 2. (e) Order 3.

Figure 3.3: Plots in the s-t plane (scales in grid points) of the fast variable x, in its unbalanced and balanced forms. As expected, increasing the order of the expansion leads to a solution that more clearly exposes its unbalanced dynamics, as the slow pendulum-like motion is filtered out to a larger extent.

this calculation being continued to arbitrary order, since

$$x^N \propto s^N \mathrm{e}^{-b|s|} \text{ as } |s| \to \infty, \tag{3.73}$$

but since the asymptotic series diverges we do not consider any higher terms than those already derived. In any event, the series only remains valid up to $\mathcal{O}(1/\epsilon)$.

The balance relations derived so far may be evaluated by comparing $x(s,t)$ with $x(s,t) - x_{\mathrm{bal}(s,t)}$, for various orders of the expansion; the second quantity should show the absence of slow motion, and should consist largely of fast waves, which will become clearer as n is increased. The plots in figure 3.3 show sequentially the full solution, $x - x_{\mathrm{bal}}$ for $n = 0, 1, 2, 3$. The model parameters were chosen as $a = 2$, $b = 1$ and $\epsilon = 0.125$. There were 250 grid points, and the solution was run over 1000 time steps with $\Delta s = 0.1004$ and $\Delta t = 1/2\epsilon b \Delta s = 0.0063$. The initial conditions were all zero, and a random forcing was applied, with $A = 0.01$, $a_{\mathrm{f}} = 10$, $b_{\mathrm{f}} = 1$, and a damping strength of $\delta = 0.01$. The results shown are qualitative, resulting as they do from different realisations of the random forcing, but they illustrate the effectiveness of the balancing procedure. Such balance relations are extremely useful for diagnosing the wave activity, and in the next section we use the balanced solutions to isolate the energy associated with the fast waves.

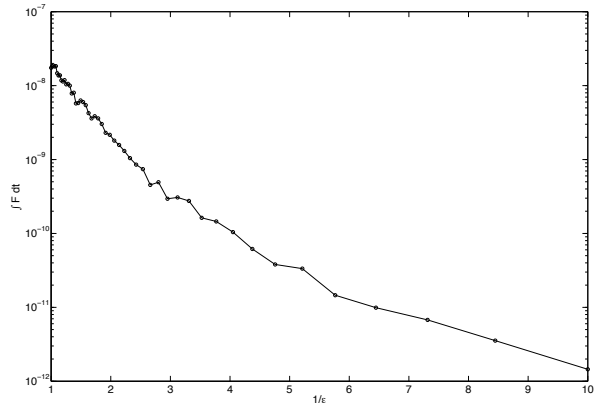


Figure 3.4: Time integrated wave flux for varying ϵ.

3.2 Results

In this section we present numerical results for the forced eLK model. For every experiment, we took a domain with $L = 25$ using 250 grid points, and set $\Delta t = 0.25\epsilon b \Delta s$. The system was started from rest with zero initial conditions in each experiment, and forcing amplitudes were chosen such that $|\theta(t)| < \pi$. For convenience, parameters for each numbered experiment are included in table 3.2 at the end of the section.

3.2.1 Varying ϵ

The exponential smallness of Rossby wave may be tested by measuring the time integrated flux of fast waves. The Rossby number ϵ was varied throughout the range $[0.1, 1]$, and the other parameters for this experiment (experiment 1) were $A = 0.01, \delta = 0.1, a_{\mathrm{f}} = 20, b_{\mathrm{f}} = 1, a = 2, b = 1, N_t = 2 \times 10^5, l = 25, \Delta s = 0.0502$, and $\Delta t = 0.1004$. The system was allowed to reach statistical equilibrium, and $\int F(l,t)dt$ was calculated over the final 10^4 time steps. In figure 3.4 the resulting flux is plotted against $1/\epsilon$; the approximately straight line dependence shows that it is indeed exponentially small in ϵ.

We next attempt to correlate the flux with the power spectral energy at the $1/\epsilon$ frequency (experiment 1a). This is the frequency at which the fast waves should be most effectively excited by the slow motion, although we still expect the excitation to be weak. For these experiments we took $A = 0.1, \delta = 0.2, a_{\mathrm{f}} = 0.2$, and $\epsilon \in [0.5, 1]$. Long runs were required to produce reasonable results; we took $N_t = 1.1 \times 10^6$, and averaged over the final 10^6 time steps. Figure 3.5 shows the power spectra that were obtained (using a Hamming window of 501 points), which all possess a characteristic kink around the resonant frequency of $1/\epsilon$. The correlation between the energy at this frequency and the flux is shown by figure 3.6. This correlation is generally quite good; it would be interesting to see whether this result generalises to a more physically realistic problem.

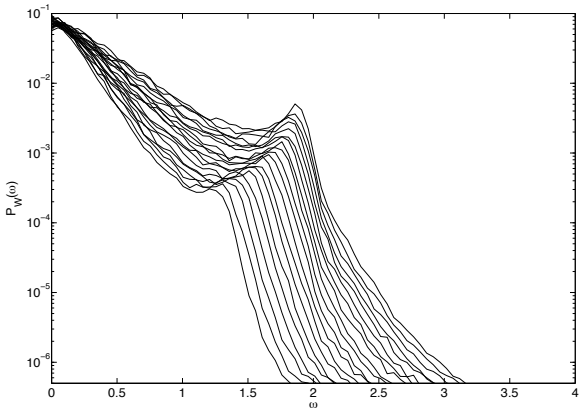


Figure 3.5: Power spectra obtained by varying the Rossby number ϵ.

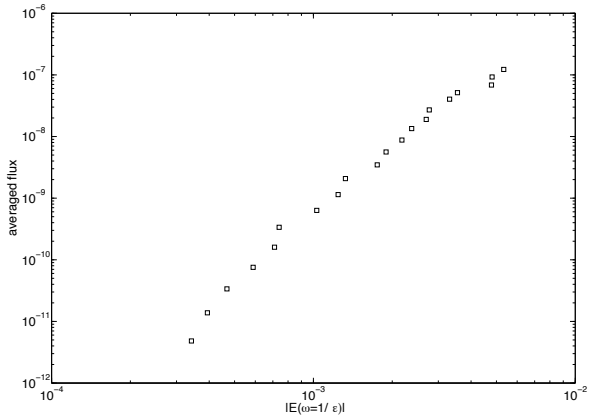


Figure 3.6: Energy at the resonant frequency against flux for varying Rossby number.

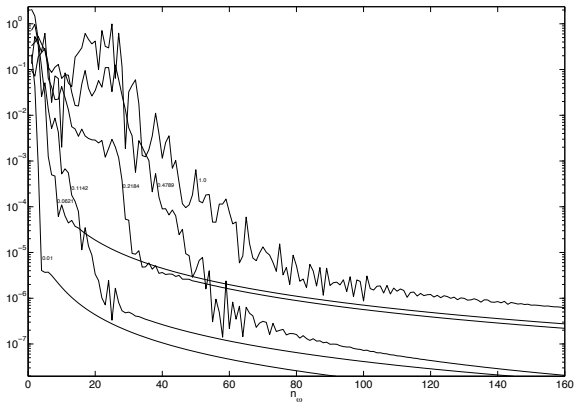


Figure 3.7: The pendulum power spectrum for varying a_{f} (marked against the relevant line), obtained by Welch periodogram with two Hamming windows of length 3000.

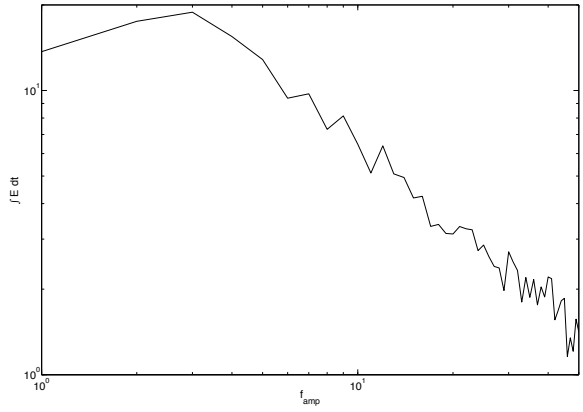


Figure 3.8: Pendulum energy against inverse correlation time, a_{f}.

3.2.2 Varying inverse correlation time, $a_{\mathbf{f}}$

In experiment 2, we first examine the power spectrum of the slow pendulum $\mathcal{P}(\theta(t))$, which like that of the underlying forcing, is expected to have an exponential form. Figure 3.7 plots the low frequency end of the power spectrum for $A = 0.1, \delta = 0.2, a_{\mathrm{f}} \in [0.01, 1], N_t = 7 \times 10^5$, and $\epsilon = 1/8$ with other parameters as before. The system was allowed to equilibrate by examining only the final tenth of the data points.

The next two figures show how the time-integrated energy and flux depend on the inverse correlation time, a_{f}. Figure 3.8 shows time-integrated energy from experiment 3 using $A = 0.1, \delta = 0.2, N_t = 10^5$ and $\epsilon = 1/8$, with $a_{\mathrm{f}} \in [0, 50]$. Figure 3.9 shows the time integrated-flux for experiment 4 with $a_{\mathrm{f}} \in [0, 5]$ and $N_t = 5 \times 10^4$, with other parameters the same.

Next we attempt to correlate the power spectral energy at the frequency $\omega = 1/\epsilon$ with the time averaged flux. Given that the power spectrum is not particularly well resolved in figure 3.7, much longer simulations are required. In experiment 5, we take $A = 0.1, \delta = 0.2, \epsilon =$

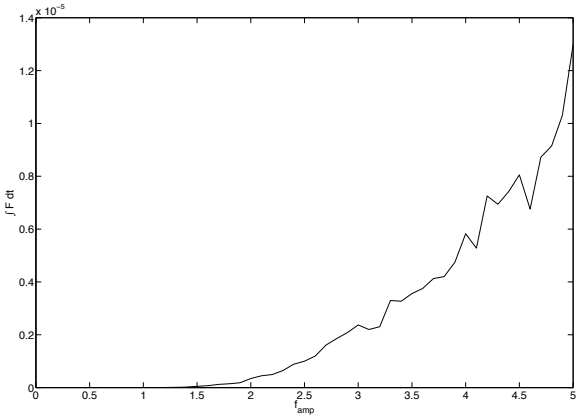


Figure 3.9: Wave flux against inverse correlation time, a_{f}.

$1/8, a_{\mathrm{f}} = 0.1, 0.2, 0.5, 1, 2, 5, 10, N_t = 1.1 \times 10^6$ and allow the system to equilibrate for 10^5 time steps. Figure 3.10a shows that the resulting power spectra are much clearer than before. The Welch estimated power spectra obtained using a 1000 point Hamming window were measured at $\omega = 1/\epsilon = 8$, and these results are plotted against the time averaged flux in figure 3.10b. This experiment was repeated using a larger number of values of a_{f} (experiment 6), and the results are shown in figures 3.10c and 3.10d. In experiment 7 we repeated the simulation with $\epsilon = 1/5$; the results are the subject of figures 3.10e and 3.10f, which both demonstrate good correlation.

3.2.3 Wave radiation and unbalanced motion

Wave radiation allows the system to shed excess energy, and even with no damping the energy can be bounded providing the forcing is not too strong. A statistical equilibrium between forcing and dissipation is reached. Figure 3.11 shows the results of experiment 8 with $A = 10^{-6}, \delta = 0, a_{\mathrm{f}} = 1, N_t = 4 \times 10^6$ and $\epsilon = 1/8$, with other parameters as before. Figures 3.12 and 3.13 show results for varying A, with $A \in [0, 1], N_t = 10^4, \delta = 2A$, and $\epsilon = 1/8$, with other parameters as before (experiment 9).

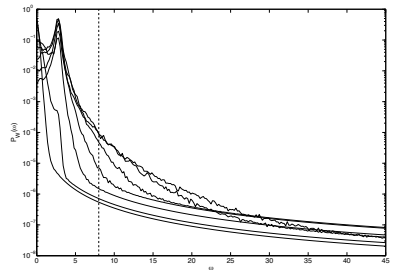
Having developed a series of balance relations for the system, we now use them to isolate the fast waves from the balanced motion, in an effort to attempt to find a correlation between the energy associated with the unbalanced motion and the flux. To this end, we compute the wave energy E^{w}

$$E^{\mathrm{w}} = \int_0^L {x_{\mathrm{w}}}_s^2 + {x_{\mathrm{w}}}^2/2 + {y_{\mathrm{w}}}^2 \; ds, \tag{3.74}$$

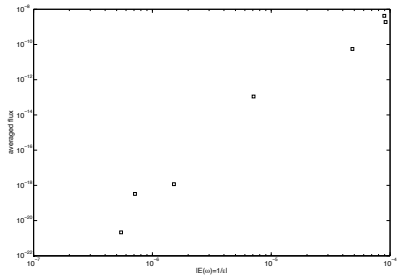
where

$$x_{\mathrm{w}} = x - x_{\mathrm{bal}}, \quad y_{\mathrm{w}} = y - y_{\mathrm{bal}}. \tag{3.75}$$

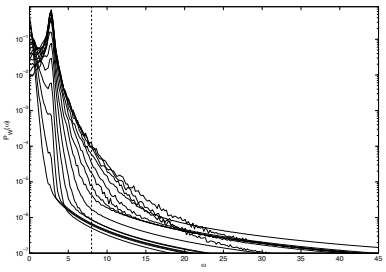
Figure 3.14 shows a typical plot of the integrand quantity for the same parameters used to generate figure 3.3. In the final two experiments (experiments 10 and 11), we took $A = 0.1, \delta = 0.2, a_{\mathrm{f}} = 1, 5, \epsilon \in [0.1, 1]$, and $N_t = 10^5$. Figure 3.15 plots the time integrated flux against the



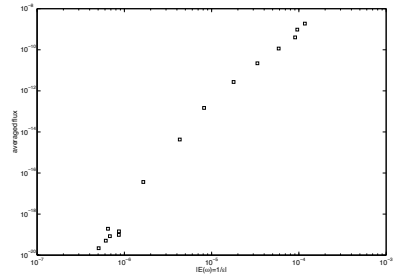
(a) Welch estimated power spectra for experiment 5.



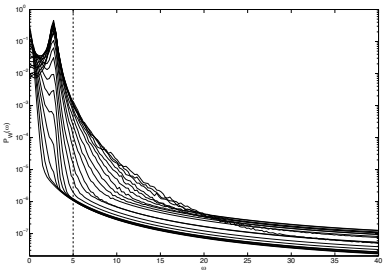
(b) Time averaged flux against $P_W(1/\epsilon)$.



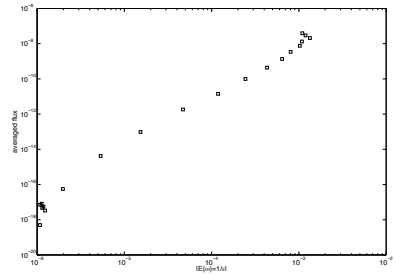
(c) Welch estimated power spectra for experiment 6.



(d) Time averaged flux against $P_W(1/\epsilon)$.



(e) Welch estimated power spectra for experiment 7.



(f) Time averaged flux against $P_W(1/\epsilon)$.

Figure 3.10: The results of experiments 5–7. The power spectra are plotted together with lines at $\omega = 1/\epsilon$.

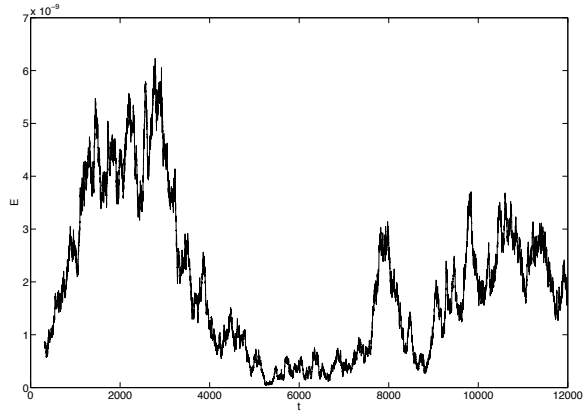


Figure 3.11: Pendulum energy over long times, in the absence of damping (experiment 8).

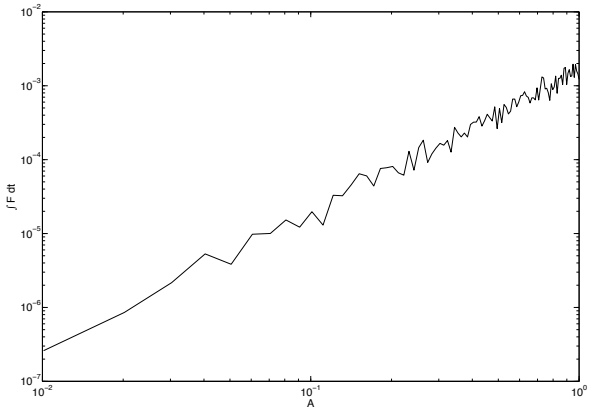


Figure 3.12: Wave flux against forcing amplitude (experiment 9).

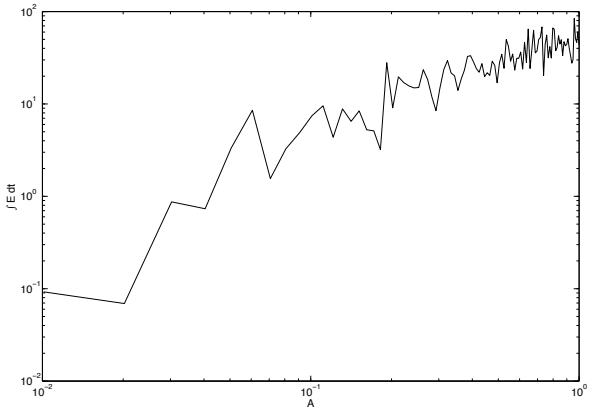


Figure 3.13: Pendulum energy against forcing amplitude (experiment 9).

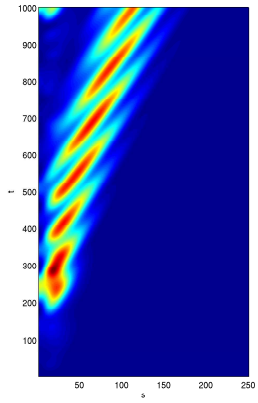


Figure 3.14: The quantity $x_{\mathrm{w}_\mathrm{s}}^2 + {x_\mathrm{w}}^2/2 + {y_\mathrm{w}}^2$ for $A = 0.01, \delta = 0.02, \epsilon = 1/8, a_\mathrm{f} = 10$.

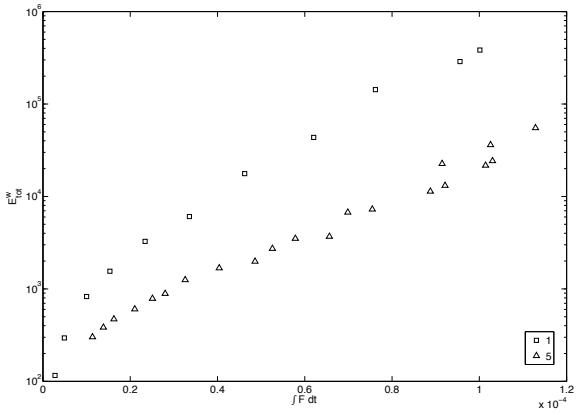


Figure 3.15: Fast wave total energy plotted against total flux, for various values of ϵ, with $a_\mathrm{f} = 1$ (squares) and 5 (triangles).

time integral of E^w for both values of a_f. The plots show that the flux correlates well with this "unbalanced energy", which shows that the balance relations are a success when it comes to isolating the unbalanced part of the motion even in this scenario where there is forcing of the fast variables. Combining this with the spectrum at $1/\epsilon$ allows us to predict the energy in the unbalanced part of the solution successfully.

From these results we can draw two general conclusions. First, the correlation between the pendulum energy at the $1/\epsilon$ frequency is strong; such a simple criterion may well be relevant to other physical problems, where power spectra are easily measured (i.e. the ocean). Second, the flux also correlates well with the energy of the fast, unbalanced dynamics, and hence the power spectral energy is a useful tool for diagnosing fast wave energy.

3.2.4 The forced shallow water equations

We may model forcing of the ocean by surface winds by adding a smooth, spatially varying random forcing to the velocity fields in rotating shallow water equations. We use a Gaussian

Experiment	A	δ	ϵ	a_{f}	N_t	N_d
1	0.01	0.1	$[0.1, 1]$	20	2×10^5	10^4
1a	0.1	0.2	$[0.5, 1]$	5	1.1×10^6	10^6
2	0.1	0.2	$1/8$	$[0.01, 1]$	7×10^5	7×10^4
4	0.1	0.2	$1/8$	$[0.01, 50]$	10^5	–
4	0.1	0.2	$1/8$	$[0.01, 5]$	5×10^4	–
5	0.1	0.2	$1/8$	$[0.01, 10]$	1.1×10^6	10^6
6	0.1	0.2	$1/8$	$[0.01, 5]$	1.1×10^6	10^6
7	0.1	0.2	$1/5$	$[0.01, 5]$	1.1×10^6	10^6
8	10^{-6}	0	$1/8$	1	4×10^6	–
9	$[0, 1]$	$2A$	$1/8$	1	10^4	–
10	0.1	0.2	$[0.5, 1]$	1	1.1×10^5	10^5
11	0.1	0.2	$[0.5, 1]$	5	1.1×10^5	10^5

Table 3.2: Experimental parameters for this section.

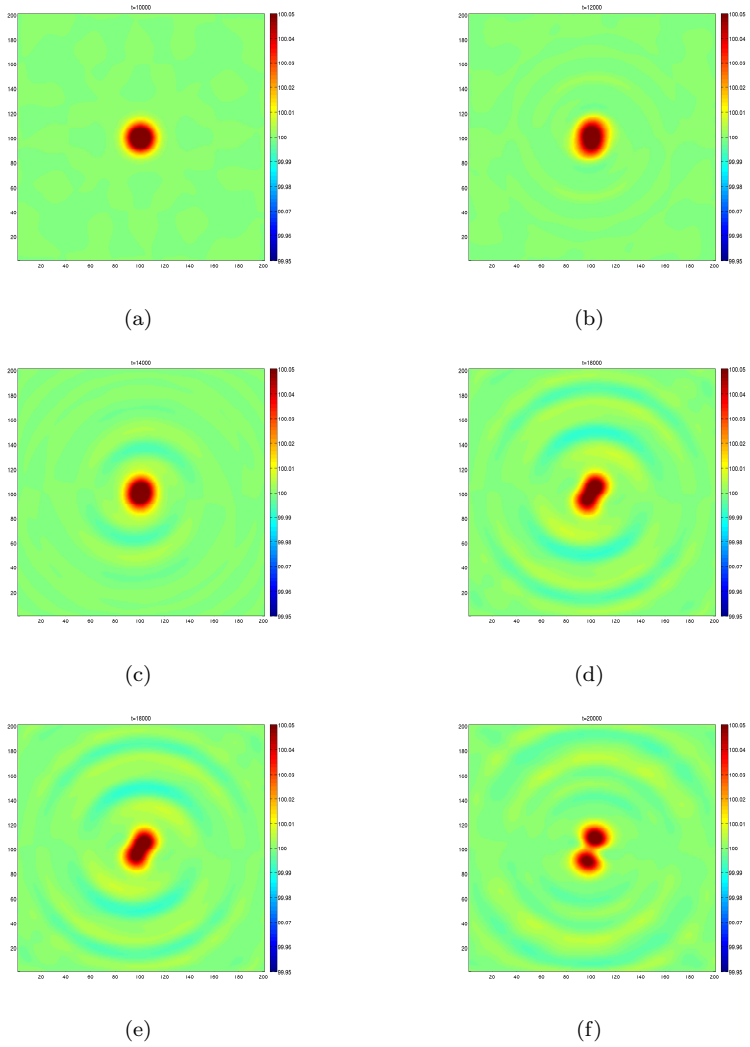
"patch" of forcing in the centre of the domain, and introduce a smooth change of sign in order not to unnecessarily add momentum to the system. For initial conditions, we take a steady, axisymmetric vortex with a Gaussian PV profile at the centre of the domain. The forcing vector is determined by two independent, smooth random variables, that have the same correlation function as 3.34. Calculation of these initial conditions is carried out in Appendix B. We solve the shallow water equations on a staggered grid using a third-order Adams-Bashforth scheme. The boundary of the domain features sponge layers for absorbing outgoing waves, though no damping is applied to the continuity equation in order to conserve mass. Details of the solver are given in Appendix C, and of the sponge layers in Appendix F. In the experiments, we vary f in the range 0.001 – 0.007, and measure the total outgoing wave energy flux just inside the boundaries.

For these experiments we take a grid of 500×500 grid points, and a sponge 10 grid points in thickness using a shifted hyperbolic damping profile; $\Delta t = 10, \Delta x = \Delta y = 5 \times 10^4, g = 9.81, H = 100$, and the experiment is allowed to run for 10^4 time steps. For the forcing, we take an amplitude of $A = 0.00001$, $a_{\mathrm{f}} = 0.1$, and use a "patch" of the form $\exp(-r^2/R^2)$ where $R = 5 \times 10^5$. Various sponge layer profiles were evaluated (see Appendix F) and the shifted hyperbolic profile employed here was shown to give the best performance for this particular scheme and sponge layer thickness.

Figure 3.16 shows several snapshots of the height field for a typical run with $f = 0.003$, and figure 3.17 shows the corresponding flux that is measured. The resulting correlation between f and the total flux over time is shown in figure 3.18. It seems likely that much longer runs, and/or averaging over many runs for each value of f is required to obtain improved results.

3.3 Conclusions

In this chapter we have investigated wave radiation as a mechanism for the maintenance of balance in a simple model that has relevance to geophysical flows. In the small ϵ limit, corresponding to small Rossby number in the fluid analogy, the frequency gap between balanced motion and fast waves means that the waves are only weakly excited, and numerical results



(a) (b)

(c) (d)

(e) (f)

Figure 3.16: Snapshots of the height field for the forced vortex ($f = 0.003$); the time step is marked in each case. The vortex sheds spiral I-G waves as it is continually deformed, and eventually fissions.

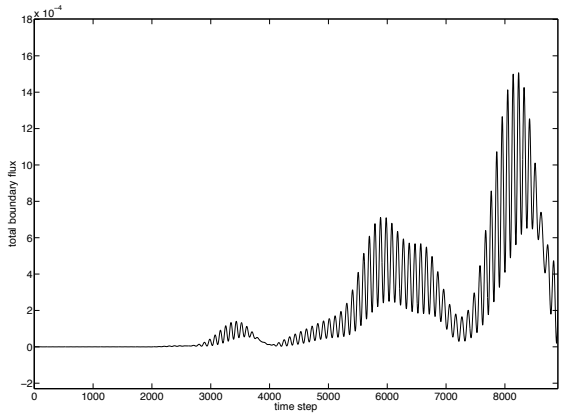


Figure 3.17: The flux over time obtained using the 2-d forced shallow water model for $f = 0.003$.

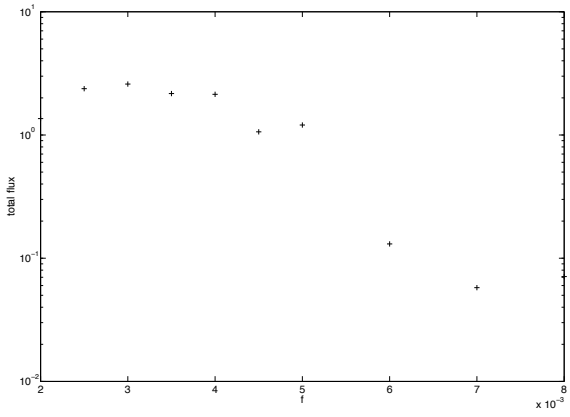


Figure 3.18: Total flux over time for varying f.

show that these waves are exponentially weak in the Rossby number. This new finding has implications for the real ocean, where spontaneous generation has been proposed as a candidate mechanism for the downscale transfer of energy. The exponential scaling demonstrated here would suggest that for Rossby numbers typically encountered in the ocean, the effect is far to weak to account for a significant part of this transfer. This agrees with recent high resolution simulations by Danioux et al [13] of spontaneous inertia-gravity waves generation by surface-intensified, near-balanced motion. They found that such spontaneously generated waves had energies orders of magnitude lower than I-G waves in the ocean, lending evidence to the notion that spontaneous generation is not a significant mechanism.

By examining the power spectra of the slow variables, we showed how it is possible in such systems to infer some basic fast motion quantities such as the flux. By deriving higher order balance relations for the slow motion, the fast motion was separated from the balanced motion, allowing the energy of each part to be extracted. The correlation between this wave energy in the interior and the flux at the boundary may go some way to explaining how close to the slow manifold the trajectories remain, as a result of wave radiation.

The model we have used is an extremely useful one because the dissipation is applied only to the fast variables, whereas damping terms are less discriminate in the variables that they affect. Such a model holds promise for further work, perhaps by introducing chaos into the balanced dynamics as suggested in [79].

Chapter 4

Wave damping and quasimodes

4.1 Quasimodes

4.1.1 Motivation

Having studied a system which maintains a statistically steady total energy by wave radiation, we now consider a mechanism by which mode-like perturbations in 2-d fluids decay; specifically, we study *quasimodes*. Such solutions arise for conservative systems, including several fluid systems of interest. Unfortunately, for physically accurate systems it is often very difficult or even impossible to derive analytical results on their Landau poles. We would be interested in studying quasimodes in the RSWEs, however this is an example of a system for which a study of quasimodes requires numerical methods. Rather than considering "real" fluid models and having to resort to numerical techniques, we instead take a toy model as the starting point. This system permits results to be found analytically via an asymptotic approach, is simple to simulate computationally, and also serves as a useful, mathematically uncomplicated example of Landau poles.

In this chapter, Landau poles are first introduced and defined, and it is shown how they arise in a fluid system using the well worn example of a perturbed axisymmetric vortex. A novel toy model (an extension of a model originally due to Lamb) is presented that exhibits quasimodes that may be revealed by hand. These quasimodes are explained mathematically and estimates for their associated decay rates are derived, which are then checked numerically. By doing so, we gain (rather less painfully) some understanding of quasimodes in 2-d inviscid fluids and the shallow water equations.

4.1.2 What are Landau poles?

Consider a conservative linear system whose solution can be found using integral transforms. The nature of the computed transform reveals the structure of the solution, even before inversion by contour integration, i.e. the spectrum of the resolvent operator may possess poles and branch cuts that contribute to the solution. For example, let us take such a solution $x(t)$ that may be written in terms of its Laplace transform $\tilde{x}(\sigma)$ as

$$x(t) = \oint \tilde{x}(\sigma) \mathrm{e}^{\sigma t} d\sigma, \tag{4.1}$$

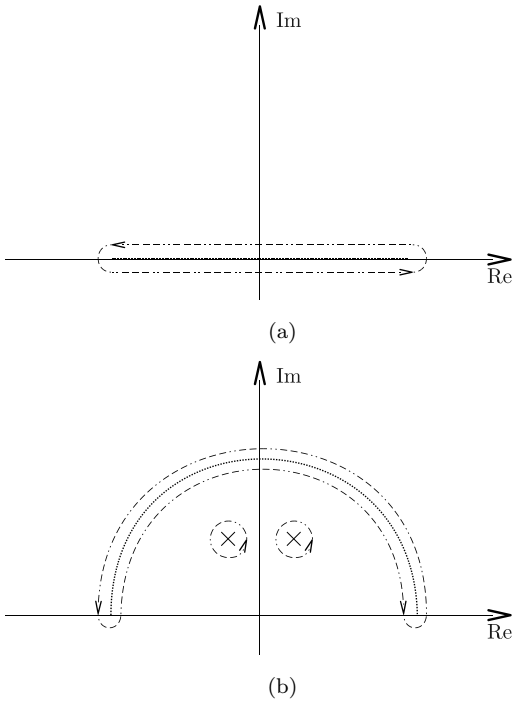


Figure 4.1: Panel a shows a branch cut in $D(\sigma)$ (dotted line) and the corresponding integration contour. In panel b, the branch cut and contour have been deformed, and in doing so two poles have been uncovered, which contribute to the integral.

where

$$\tilde{x}(\sigma) = \frac{1}{D(\sigma)} f(\sigma) \tag{4.2}$$

and the function $D(\sigma)$ is termed the *dispersion relation*. $D(\sigma)$ will dictate the nature of the solution: if it is multivalued, then the inversion integral will feature branch cuts, corresponding to parts of the continuous spectrum. If $D(\sigma)$ has poles, these correspond to eigenvalues of the system, i.e. eigenmodes. Now, we can **analytically continue the function $D(\sigma)$ in the complex plane**, and in doing so involve poles that lie on a **different Riemann sheet** in the contour integration procedure [1, 10, 11, 12]. Such poles are the **Landau poles** of the system. Such a deformation procedure is illustrated in figure 4.1. They do not correspond to a mode of the system, but nevertheless contribute to behaviour which is mode-like.

Figure 4.2 illustrates the idea of poles lying on another Riemann sheet, which may be uncovered by rotating a branch cut. Note that by deforming contours and uncovering Landau poles we gain nothing; their contribution is part of the continuous spectrum of the undeformed problem, and we simply make the solution more informative. By rotating the branch cut, the poles that were previously on the "wrong" Riemann sheet, are now uncovered and feature in the contour integral.

Each Landau pole is associated with a quasimode, and a system which possesses decaying quasimodes is said to be affected by *Landau damping*. We next consider some systems that exhibit Landau damping, where this damping is accompanied by a simple physical interpretation.

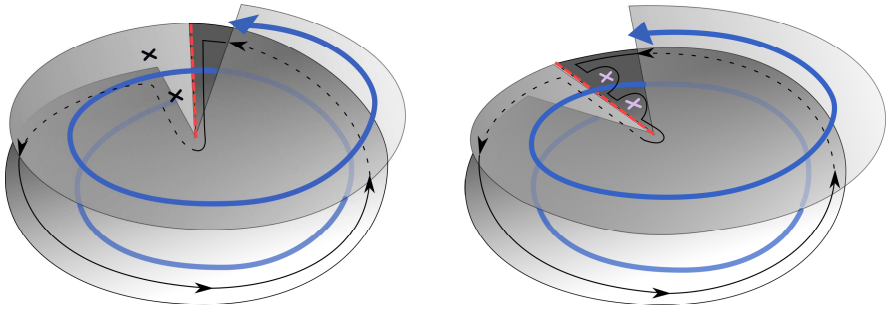


Figure 4.2: Schematic of contour integrals involving poles and branch cuts. The two Riemann sheets are illustrated, and the direction of increasing angle is shown by the blue arrow. There is a single branch cut (red dashed line) which lies along the entire positive imaginary axis. Contours of integration — which here are large circles in the complex plane — are shown by the black lines with arrows; the lines are dashed when they are obscured by the upper sheet. In the left panel there is a Landau pole on the "wrong" Riemann sheet (shown with a cross). In the right panel, the branch cut is rotated, and now the pole happens to lie on the "correct" Riemann sheet, where it may be "picked up" during integration with an appropriately deformed contour which encircles the poles.

4.2 Landau poles in fluid systems: quasimodes and Landau damping

In this section, we gradually build up a series of models. The first supports modes in the form of Rossby-Kelvin waves, in the second by adding rotation and nonzero Froude number we obtain a radiation-induced instability, and finally, by perturbing the PV profile, obtain a system with quasimodes.

4.2.1 Rossby-Kelvin modes

The two main dimensionless quantities that will be used to characterise flows under consideration will be the Rossby number, describing the balance between inertial and rotational effects on the dynamics, and the Froude number (a geophysical equivalent of the Mach number), which gives an indication of the role of compressibility. They are defined as

$$R_o = U/fL \quad \text{and} \quad F_r = U/\sqrt{gH}, \tag{4.3}$$

where U, L and H are respectively a characteristic velocity, length scale and thickness for the flow. Let us consider a vortex that is almost incompressible, i.e. that satisfies

$$\delta\rho/\rho_0 \ll 1. \tag{4.4}$$

Now

$$\frac{dp}{dt} = \left(\frac{dp}{d\rho}\right)_\eta \frac{d\rho}{dt} = c_s^2 \frac{d\rho}{dt}, \tag{4.5}$$

where c_s is the sound speed of a compressible fluid. Hence we may write

$$\delta p \sim c_s^2 \delta\rho. \tag{4.6}$$

Now consider the dimensional balance between the acceleration and pressure gradient terms in the Euler equations:

$$\frac{U^2}{L} \sim \frac{1}{L}\frac{\delta p}{\rho_0}; \tag{4.7}$$

combining this with the approximation in (4.6) gives

$$U^2 \sim \frac{c_s^2 \delta\rho}{\rho_0}. \tag{4.8}$$

Hence

$$\frac{\delta\rho}{\rho_0} \ll 1 \Leftrightarrow \frac{U^2}{c_s^2} \ll 1. \tag{4.9}$$

Hence, we may associate the incompressible approximation with $F_r \ll 1$.

Let us take the incompressible limit (achieved as $F_r \to 0$) without rotation ($f = 0$), together with a vortex with an axisymmetric basic state governed by a very simple radial vorticity distribution $\xi(r)$. Such a problem — framed in cylindrical coordinates — is considered by Balmforth et al [4]:

$$\xi(r) = \begin{cases} 1, & r \leq 1, \\ 0, & r > 1. \end{cases} \tag{4.10}$$

Such a vortex allows the existence of a population of Rossby-Kelvin modes. These disturbances move benignly around the vortex edge. For the perturbations, if we introduce vorticity ξ and streamfunction ψ, and for the basic flow a vorticity Z and angular rotation rate Ω, then the disturbances to the constant state are fully described by

$$\partial_t \xi + \Omega \partial_\theta \xi - \tfrac{1}{r}\psi_\theta Z' = 0. \tag{4.11}$$

Here $\Omega = V/r$, where V is an azimuthal velocity, and $Z = r\Omega' + 2\Omega$. Dashes denote differentiation with respect to r. The perturbation vorticity is linked to the streamfunction by

$$\xi = \hat{\mathbf{z}} \cdot (\nabla \times (u, v, 0)) = v/r + v_r - u_\theta/r = \psi_{rr} + \psi_r/r + \psi_{\theta\theta}/r^2. \tag{4.12}$$

To proceed and solve for ψ in order to determine the dispersion relation for the Rossby-Kelvin modes, we Fourier transform (4.12), after first determining the form of the vorticity perturbation, ξ. Now $Z(r) = H(1 - r)$ and so $Z' = -\delta(1 - r)$; substitution in (4.11) demands that

$$\xi = \mathrm{e}^{\mathrm{i}(m\theta - \omega t)}\delta(r - 1). \tag{4.13}$$

Hence, for $r \neq 1$, we may write the Fourier transform of (4.12) as

$$\hat{\psi}_{rr} + \tfrac{1}{r}\hat{\psi}_r - m^2/r^2\hat{\psi} = 0, \tag{4.14}$$

with solutions of the form

$$\hat{\psi} = \mathrm{e}^{\mathrm{i}(m\theta - \omega t)}\hat{\phi}(r), \tag{4.15}$$

with

$$\hat{\phi} = \begin{cases} Ar^m & r < 1, \\ Ar^{-m} & r > 1. \end{cases} \tag{4.16}$$

Here the condition of decay towards infinity has been imposed. Now

$$[\hat{\psi}]_{r=1^-}^{r=1^+} = 1 = -2mA, \tag{4.17}$$

giving $A = -\frac{1}{2m}$. Finally, substituting this result back into (4.11) gives

$$\begin{aligned} 0 &= [-\mathrm{i}\omega + \mathrm{i}m\Omega(r) - \mathrm{i}m/2mr]\delta(r-1), & (4.18)\\ \Rightarrow 0 &= -\omega + m/2 - 1/2, & (4.19)\\ \Rightarrow \omega &= m/2 - 1/2. & (4.20) \end{aligned}$$

This simple expression is the **Rossby-Kelvin dispersion relation**.

4.2.2 Adding a radiation induced instability

By very slight modification of the basic assumptions used to derive (4.20), it is possible to excite a weak mode of instability in the flow as studied by Broadbent & Moore [7], Kop'ev & Leont'ev (for the irrotational case) [49], Ford et al [25], and Ford alone [24] (the latter paper being the most relevant to the current work). By assuming $F_r \ll 1$, Ford showed that ω acquires a small imaginary part. Such weak growth is actually found to be on the order of F_r^4. To see this, note that by taking F_r and f nonzero, the streamfunction requires modification from its previous form (r^{-m}) for the outer wave region:

$$\psi = A[r^{-m} + \mathrm{i}C(r^m - r^{-m})], \tag{4.21}$$

where

$$C = \frac{\pi m F_r^{2m}}{2(m!)^2}\left(\frac{\omega_0^2 - f^2}{4}\right)^m \left(\frac{\omega_0 - f}{\omega_0 + f}\right) + \mathcal{O}(F_r^{2m+2}). \tag{4.22}$$

Substitution into the first part of (4.17) gives

$$A = -(1 - \mathrm{i}C)/2m, \tag{4.23}$$

and hence from (4.11) we obtain the new dispersion relation for these unstable modes:

$$\omega = 1/2(m - 1 + \mathrm{i}C). \tag{4.24}$$

Since $C \sim F_r^{2m}$, the $m = 2$ mode — the first nontrivial mode, since $m = 1$ represents a plane translation rather than a distortion of the vortex — is also the one that grows the most aggressively, and is the most easily observed. All modes are unstable.

4.2.3 Critical layers, Landau damping, and quasimodes

Although regularly seen in the context of plasma physics in which it first arose, the process of Landau damping (in plasmas a wave-particle resonant interaction) is also relevant to certain fluid flow problems. In §4.2.1 we introduced a vortex with a discontinuous vorticity profile, and then examined an instability resulting from the addition of compressibility. Let us now consider the aforementioned competing effect due to Landau damping, which may be introduced simply by smoothing out the PV profile from its hard step form — a procedure considered in great detail by Balmforth et al [4]. Qualitatively, the Rossby-Kelvin modes are replaced by a decaying

disturbance, which mimics a Rossby-Kelvin mode and decays exponentially. In fact, the vortex with a smooth PV profile does not support "real" modes, and the observed disturbances are examples of *quasimodes*, which are a cooperative effect of the continuous spectrum. Balmforth et al studied the appearance of quasimodes as the PV step is smeared out to larger radii, by considering a thin PV "skirt" placed around a localised vortex. Dynamically, the addition of such a skirt allows the existence of a *critical radius*; a distance at which the frequency of a mode supported by the vortex is equal to the local rotation frequency of the fluid. If we denote the critical radius by r_c, then by definition

$$m\Omega(r_c) = \omega. \tag{4.25}$$

If the dispersion relation for the vortex waves is known, then the critical radius may be computed, e.g. for an $m = 2$ Rossby-Kelvin wave of §4.2.1 we find $r_c = \sqrt{2}$. Qualitatively, when the PV distribution is smoothed out so that it reaches the critical radius, there is a thin annular region of activity or *critical layer* in which the vorticity distribution becomes wrapped up tightly into a spiral, without decaying. The associated quasimode decays exponentially, and the vortex relaxes to axisymmetry.

Mathematically, quasimodes may be revealed in more that one way. In a classic paper by Briggs et al [6], an energetic (integral) argument was used to uncover quasimodes in an incompressible, inviscid 2-d shear flow. However, quasimodes may also be uncovered by another method (as done in [6] to complement the integral derivation), by considering the spectrum of the operator describing the evolution of the system. Systems possessing quasimodes have a spectrum that is multivalued in the complex plane, and hence involves branch cuts which must be chosen carefully. The absence of true eigenmodes implies that there are no poles in the complex plane; any behaviour akin to a mode is due to contributions from the continuous spectrum, i.e. the branch cut. However, such mode-like behaviour might be more clearly uncovered by noticing that *there may well be poles lying on another Riemann sheet.* There is nothing gained by bringing these poles into play — their contributions are included in the continuous spectrum — but it is useful to **rearrange the branch cuts so that the poles are made more visible in the spectrum**, and their contribution easily seen. In some cases it is not possible to derive analytic expressions for the locations of such poles, so a numerical method must be used to hunt for them, as was done for example in [69].

The system studied by Schecter at al in [69] features Vortex Rossby waves that may on the one hand be destabilised by the loss of energy through radiation of spiral inertia-gravity waves (IGWs), but on the other hand weakened by Landau damping. They studied a vortex with a hyperbolic PV profile, with smoothness controlled by a parameter Δ. $\Delta = 0$ corresponds to a hard step profile, becoming increasingly smooth as Δ is increased. It was shown that the winning mechanism depends on this parameter. The respective growth/decay rate of each effect increases in magnitude as $\Delta \to \infty$. For $\Delta = 0$, there is no Landau damping, and so radiative instability wins out. However, increasing Δ sees the critical layer damping effect overtake radiative pumping, and so sufficiently smooth vortices always axisymmetrize. The transition between growth and decay is seen at $\Delta \approx 0.2$ in their model, for a particular choice of parameters. These results are reproduced using a 1-d linear solver; details may be found in appendix B.

Briggs et al [6] studied its role in the case of incompressible, inviscid shear flow: they considered a piecewise constant axisymmetric initial vorticity profile — i.e. a step in the radial direction — and were able to show using energetic arguments that the existence of even the slightest gradient in the profile was sufficient to exclude the presence of unstable modes. In an equivalent analysis they showed that the disappearance of the unstable mode is due to Landau damping, by analytically continuing the transformed solution, and determining the placement of the Landau poles. It is important to remember that any continuation procedure does nothing to the content of the solution, indeed, Turner & Gilbert [77] point out, any damping is merely a cooperative effect of the continuous spectrum — we merely reveal it more clearly it by analytic continuation.

In order to reduce the complexity of a study on Landau damping and more closely focus on the relevant points, we take a step back from "real" fluid or plasma systems, and instead work with a slightly altered version of a toy model originally due to Lamb [50]. This model system is not dissimilar in many ways from the extended Lorenz Krishnamurthy model; for example, it features an oscillator coupled to a dispersive string, which allows wave radiation to infinity. The model has subsequently been used as a basis for understanding a variety of physical problems [33, 5]. The model departs from geophysical reality in that the Landau damping is entirely due to the presence of *wave radiation* in the model (implemented numerically using the exact DNRBC from §2.3.3 of chapter 2), but despite this major difference it is still a rich and interesting model. In section 4.3 we introduce Lamb's model and explain our modification, and in section 4.3.4 we show how the Landau poles affect the decay rate for the system by deforming the contours used to invert the transforms that arise in the solution. In section 4.3.5 these are compared with results from numerical experiments. Finally, in section 4.4 a small linear damping term is added to the wave equation, and the effect of doing so is considered. Details of the derivations of the spectral expansion for the system are deferred to Appendix E.

4.3 Quasimodes in an extended Lamb model

4.3.1 Lamb's original model

Lamb [50] originally introduced a simple one-dimensional model of vibrating bodies surrounded by a transmitting medium, an example of such a system being a deformable sphere suspended in a gas. Specifically, he sought to explain the appearance of apparently spurious solutions of such systems, which appear to grow exponentially with distance from the sphere. The problem surrounding the existence of these modes is easily dealt with by noting that if the vibration is started at a time $t = 0$, then disturbances with wavespeed c can at time t have travelled a maximum distance of $x = ct$, beyond which the medium is as yet undisturbed. The exponential growth in the solution actually represents the decay in the energy of the oscillating body. As an illustration of this point one may think of a bell being struck and the sound monitored by three observers at some later time: the observer nearest the bell hears the bell ringing at a reduced volume; further away, the bell sounds louder, and finally, an observer sufficiently far away hears nothing, since the wavefront has yet to reach them.

Lamb further illustrated this phenomenon by means of a toy model, which comprises a linear oscillator of mass M and frequency of oscillation ω, coupled at a point $x = 0$ to an infinite

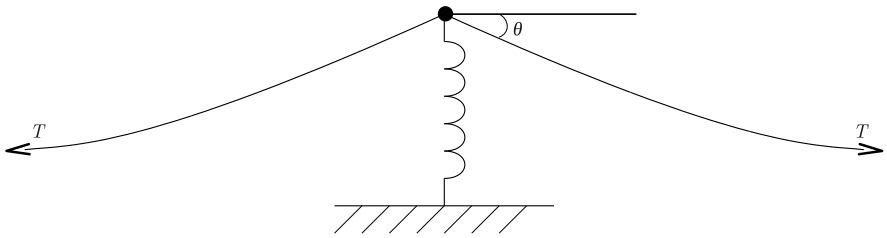


Figure 4.3: The system represented by the Lamb model, consisting of an oscillator coupled to a taut string at a point.

stretched string with a wavespeed c (fig. 4.3.1). The wavespeed is given in terms of the density and T by $c^2 = T/\rho$. If we denote the coupling constant T and the position of the string $u(x,t)$, then the model equations are

$$M\left(u_{tt}(x=0,t) + \omega^2 u(x=0,t)\right) = T[u_x]^+_-, \tag{4.26}$$

$$u_{tt} = c^2 u_{xx}. \tag{4.27}$$

where the limit is taken about the point $x = 0$. The first equation is derived by noting that the force on the oscillator due to the string is modelled by

$$T\sin\theta_+ - T\sin\theta_- \approx T\left[\frac{\partial u}{\partial x}\right]^+_-, \tag{4.28}$$

where we have taken the small angle approximation for $\sin\theta$. Let us consider a right-going solution for the case $x > 0$; we choose right going solutions only so as to represent waves travelling away from the oscillator. Assume a solution of the form

$$u(x,t) = \mathrm{e}^{\mathrm{i}(kx-\Omega t)}, \quad x > 0. \tag{4.29}$$

Together with (4.26) we find

$$M(-\Omega^2 + \omega^2) = 2\mathrm{i}kT, \tag{4.30}$$

and from (4.27) we have

$$\Omega^2 = c^2 k^2. \tag{4.31}$$

Combining (4.30) and (4.31) gives the dispersion relation for waves on the string

$$k^2 + \frac{\mathrm{i}k}{m} - \frac{\omega^2}{c} = 0, \tag{4.32}$$

where $m = M/2T$. Solutions are therefore

$$k = -\frac{\mathrm{i}}{m} \pm \sqrt{-\frac{1}{4m^2} + \frac{\omega}{c^2}}. \tag{4.33}$$

The key point is that this solution **does not correspond to a true mode of the system**. This is because in deriving it we have neglected the boundary condition that waves should decay towards infinity, which dictates $\mathrm{Im}\, k > 0$. However the zeros of (4.32) found have $\mathrm{Im}\, k < 0$, corresponding to a solution that grows away from the origin. This apparent paradox is resolved by realising that $u(x,t) = 0$ is the solution for $x > ct$, i.e. the initial disturbance has only

caused waves to propagate a finite distance, so the wave does not continue to grow as $x \to \infty$. This true solution — a travelling wave that grows in space up to a point where the string is abruptly quiescent — is quite clearly not a mode. The derived solution is in fact a quasimode.

4.3.2 A dispersive Lamb model

In the model all waves propagate with phase velocity and group velocity c. We modify this by adding an extra term of the form $m^2 u$ to the second equation, so that the string is governed by the Klein-Gordon (KG) equation and is now dispersive. The degree of dispersion is controlled by the constant m. The resulting dispersion relation is now of the form

$$\omega^2 = \sqrt{m^2 + c^2 k^2}. \tag{4.34}$$

We note out of interest that this expression has almost the same form as the dispersion relation for gravity waves in 2-d shallow water [26]. There is now a minimum frequency set by m, and hence it is possible for there to be a frequency gap between that of the oscillator and the lowest possible frequency mode of the string, in the case of $\omega < m$. This simple system has not yet received detailed attention in the literature, and in this chapter we probe it thoroughly. If we non-dimensionalise using the wave-speed and oscillator frequency, and write the position of the oscillator by q (i.e. $u(x = 0, t) = q(t)$), then our modified system is of the form

$$\begin{aligned} \ddot{q} + q &= T\,[u_x]^+_-\,, \\ u_{tt} - u_{xx} + m^2 u &= 0. \end{aligned} \tag{4.35}$$

The influence of dispersion is clearly seen if we plot the displacement of the string in time and space, as in the right panel of figure 4.4. In the simulation, the string is initially undisplaced and stationary, and the oscillator is given a starting velocity. We see that the shorter waves begin to catch up with their longer counterparts as time progresses, leaving behind longer waves (were $m = 0$ then the streaks would all be straight lines, as demonstrated by the left panel of figure 4.4).

We proceed by rewriting (4.35) as the dynamical system:

$$z_t = \mathcal{L} z,$$

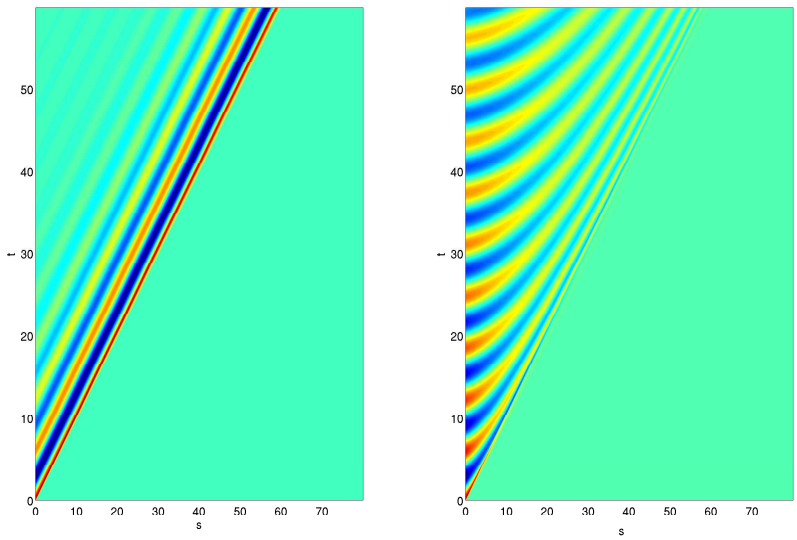
where $z = (p, q, v, u)$, $p = \dot{q}$, $v = u_t$, and

$$\mathcal{L} z = \begin{pmatrix} -q + T\,[u_x]^+_- \\ p \\ u_{xx} - m^2 u \\ v \end{pmatrix}.$$

We now derive explicit formulae for the general solution of (4.35), that is, for the operator $\exp(t\mathcal{L})$. It is useful to consider more general functions $f(\mathcal{L})$. Using Cauchy's formula, these may be written as

$$f(\mathcal{L}) = \frac{1}{2\mathrm{i}\pi} \oint f(\sigma)\,(\sigma \mathbb{I} - \mathcal{L})^{-1}\, d\sigma, \tag{4.36}$$

where the integration is over a large circle in the complex σ-plane, avoiding possible branch



(a) $m = 0$ (b) $m \neq 0$

Figure 4.4: Numerical solution of string displacement plotted in the x-t plane both without dispersion (left) and with dispersion (m = 0.8). Parameter values are $T = 0.5$, $\Delta x = 0.08$ and $\Delta t = 0.04$.

cuts. Our aim is to compute the resolvent operator $(\sigma\mathbb{I} - \mathcal{L})^{-1}$, substitute in (4.36) and reduce the integral to contributions from residues (eigenvalues of $\mathcal{L}$), and branch cuts (the continuous spectrum).

Appendix E is given over to the derivation of the spectral expansion for the model, together with the associated eigenfunctions and eigenvalues.

4.3.3 The structure of the solution

The solutions for the oscillator and the string are given in terms of their Laplace transforms in time $\tilde{p}(\sigma)$, $\tilde{q}(\sigma)$ and $\tilde{u}(\sigma, x)$ by

$$\begin{pmatrix} \tilde{p} \\ \tilde{q} \end{pmatrix} = \frac{1}{\sigma^2 + 1 + 2T\sqrt{m^2+\sigma^2}} \begin{pmatrix} \sigma & -1 - 2T\sqrt{m^2+\sigma^2} \\ 1 & \sigma \end{pmatrix} \begin{pmatrix} \tilde{w}_p \\ w_q \end{pmatrix}, \tag{4.37}$$

where

$$\tilde{w}_p = w_p + T\int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x'|}\left(\sigma w_u(x') + w_v(x')\right)\, dx',$$

and

$$\begin{aligned} \tilde{u} \;=\; & \frac{1}{D(\sigma)}\mathrm{e}^{-\sqrt{m^2+\sigma^2}|x|}(\sigma w_q + w_p) \\ & -\frac{\sigma^2+1}{2D(\sigma)\sqrt{m^2+\sigma^2}}\mathrm{e}^{-\sqrt{m^2+\sigma^2}|x|}\int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x'|}\left(\sigma w_u(x') + w_v(x')\right)\, dx' \\ & +\frac{1}{2\sqrt{m^2+\sigma^2}}\int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x-x'|}\left(\sigma w_u(x') + w_v(x')\right)\, dx'. \end{aligned} \tag{4.38}$$

Here the variables w_p, w_q etc. denote the initial conditions for each subscripted quantity, and $D(\sigma)$ is referred to as the *dispersion relation*, given by

$$D(\sigma) = \sigma^2 + 1 + 2T\sqrt{m^2 + \sigma^2} = 0. \tag{4.39}$$

It is useful to see how the solution $u(x, \sigma)$ reduces to that originally derived by Lamb for the dispersionless string, by first setting (for simplicity) $w_q = w_u = w_v = 0$, i.e. the string is initially quiescent and the oscillator sitting at the rest position at $t = 0$ is given an initial velocity. The resulting equation for the string displacement obtained by the inverse Laplace transform is

$$u(x,t) = w_p \oint \frac{e^{\sigma t}}{D(\sigma)} e^{-\sqrt{m^2+\sigma^2}|x|}\, d\sigma. \tag{4.40}$$

Setting $m = 0$ gives

$$u(x,t) = w_p \oint \frac{e^{\sigma t - \sqrt{\sigma^2}|x|}}{\sigma^2 + 1 \pm 2T\sqrt{\sigma^2}}\, d\sigma. \tag{4.41}$$

In order that the boundary condition be satisfied when $x > t$, we require $\sqrt{\sigma^2} > 0$. This choice precludes the existence of any zeros of the dispersion relation, and hence no poles are picked up when performing the integration over the large contour. Hence, for $x > t$, there is no contribution to $u(x, t)$ and we predict that the string is undisturbed beyond $x = t$, agreeing with our physical reasoning.

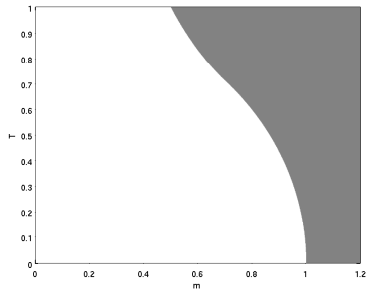
The resolvent is analytic for $\sigma \in \mathbb{C}$ except on the branch cuts $[\mathrm{i}m, \mathrm{i}\infty) \cup (-\mathrm{i}\infty, -\mathrm{i}m]$, and the possible roots of the dispersion relation, which can be interpreted as eigenvalues of $\mathcal{L}$. The dispersion relation is really the key to unravelling much of the behaviour of the system: the types of behaviour of the dynamical system found as m and T are varied are governed directly by the form of the roots of $D(\sigma)$. Care must be taken in solving (4.39), as some of the solutions

$$\sigma^* = \pm\sqrt{-1 + 2T^2 \pm 2T\sqrt{m^2 - 1 + T^2}} \tag{4.42}$$

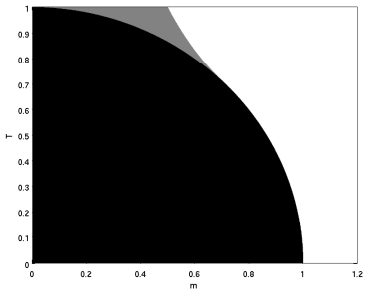
are spurious since they correspond to $\mathrm{Re}\,\sqrt{m^2 + \sigma^2} < 0$. These spurious solutions, corresponding to a different Riemann sheet from that relevant to the resolvent computation, are Landau poles.

It is now possible to characterise the different solutions σ^* that are found in various regions of the parameter space defined by m and T. The boundary that divides regions containing "true" poles and Landau poles is found by setting $m^2 + \sigma^{*2} = 0$; a quick calculation shows that the dependence on T drops out, and gives the solution $m^2 = 1$. By direct calculation it may be shown that In the region $m^2 < 1$, all the solutions are spurious as they correspond to $\sqrt{m^2 + \sigma^2} < 0$ — these are the Landau poles. For $m^2 > 1$, two of solutions become 'true' solutions associated with the positive square root, and the other two are Landau poles.

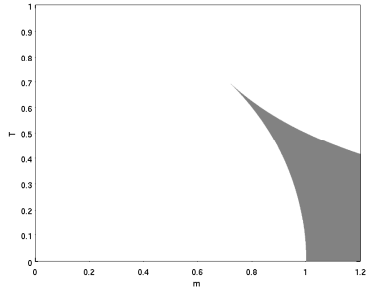
We can say more about the nature of the roots in our parameter space. If we plot the signs of $\mathrm{Re}\,\{\sigma^*(m,T)\}$ and $\mathrm{Im}\,\{\sigma^*(m,T)\}$, for $m \in [0, 1.2]$, $T \in [0, 1]$ (figure 4.5(a-d)), we find there



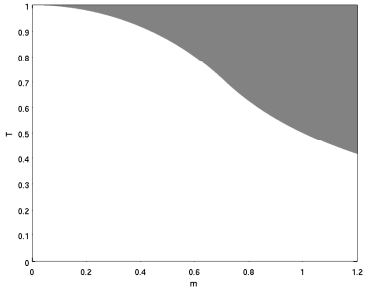
(a) Real part, negative square root



(b) Imaginary part, negative square root



(c) Real part, positive square root



(d) Imaginary part, positive square root

Figure 4.5: Plots showing the signs of the real and imaginary parts of $\sigma(m,T)$ in the m-T plane, for both possible choices of the sign of the inner the square root in the expression for σ^*. Black indicates a negative sign, white a positive sign, and grey corresponds to zero.

are four distinct regions, divided by the curves $T = 1/2m$ and $T^2 = 1 - m^2$ which coalesce at $T = 1/\sqrt{2}$ (figure 4.6). The curve $T = 1/2m$ is found by setting the contents of the outer square root in the expression for σ^* to zero and solving for T:

$$\begin{aligned} -1 + 2T^2 \pm 2T\sqrt{m^2 - 1 + T^2} &= 0 \\ m^2 - 1 + T^2 &= \left(\frac{1}{2T} - T\right)^2 \\ \Rightarrow m &= \frac{1}{2T}, \end{aligned}$$

assuming $T, m > 0$. The second curve is found by applying the same procedure to the inner square root.

The plots in figure 4.5 show the signs of the real and imaginary parts of σ^* for both the positive and negative square roots. From these plots we conclude that the types of the solutions in each region are as follows:

- $m^2 < 1 - T^2$ (region 1): there are four Landau poles, $\pm\sigma_L$, $\pm\bar{\sigma}_L$, with $\mathrm{Re}\{\sigma_L\} \neq 0$.
- $1 - T^2 < m^2 < 1/4T^2$, $T > 1/\sqrt{2}$ (region 2): there are Landau poles with $\mathrm{Re}\{\sigma_L\} \neq 0$ and $\mathrm{Im}\{\sigma_L\} = 0$.
- $1 - T^2 < m^2 < 1/4T^2$, $T < 1/\sqrt{2}$, $m > 1$ (region 3): there are four roots (of which two are eigenvalues if $m^2 > 1$) with $\mathrm{Re}\{\sigma^*\} = 0$, $\mathrm{Im}\{\sigma^*\} \neq 0$.
- $m^2 > 1/4T^2$ (region 4): there are two roots (Landau poles if $m^2 < 1$, eigenvalues if $m^2 >$

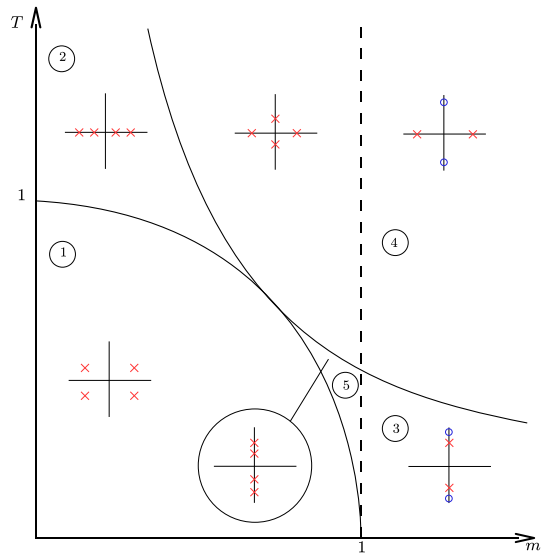


Figure 4.6: Diagram showing the positions in the complex plane of the poles of the resolvent, in each of the four regions in the m-T parameter space — these being bounded by the solid curves $T = 1/2m$ and $T^2 = 1 - m^2$. Landau poles are shown as crosses, and eigenvalues as circles.

1) with $\mathrm{Re}\{\sigma^*\} = 0$, $\mathrm{Im}\{\sigma^*\} > 0$; there are two roots with $\mathrm{Re}\{\sigma^*\} \neq 0$, $\mathrm{Im}\{\sigma^*\} = 0$.

- $1 - T^2 < m^2 < 1/4T^2$, $T < 1/\sqrt{2}$, $m < 1$ (region 5): there are four Landau poles each with $\mathrm{Im}\,\sigma^* = 0$.

This is summarised most clearly in figure 4.6.

4.3.4 Analytic continuation and hidden poles: Decay rates for the dispersive system

We now come to what is in essence the key point of this work: to show how the Landau poles, which may at first glance seem irrelevant, make a contribution to the solution, which depends on a small parameter. Recall that in section E.1, we showed that the transforms of the functions of interest have no poles in the σ plane for one branch of the square root, but possess poles for the other choice (i.e. they are on the "wrong" Riemann sheet). It would seem that these spurious poles are irrelevant to our inversion; we merely integrate around the branch cuts in order to find $q(t)$, $p(t)$ $u(x,t)$ and $v(x,t)$. However, in the case where there are no eigenmodes which would otherwise dominate, and in addition $m \ll 1$, the damping effect due to the presence of Landau poles will be evident. To reveal the contributions of these poles, we deform branch cuts and contours of integration, so that the Landau poles are "picked up" by our integral.

The aim is to determine the decay rate due to Landau damping. To measure this, we consider the energy of the oscillator over time. It is easy to show that the total energy of the system is given by

$$\frac{p^2}{2} + \frac{\omega^2 q^2}{2} + \frac{T}{2}\int_0^\infty (u_x^2 + v^2 + m^2 u^2)dx. \tag{4.43}$$

Hence, the energy of the oscillator alone can be estimated using only the first two terms in (4.43)

$$E_{\rm osc} = \frac{p^2}{2} + \frac{\omega^2 q^2}{2}. \tag{4.44}$$

We now calculate $p(t)$ and $q(t)$ using results given in Appendix E. From the spectral expansion we may write the Laplace transforms in time (denoted with a tilde) of p and q as

$$\tilde{p} = \frac{1}{D(\sigma)}\left[\left(-1 - 2T\sqrt{m^2+\sigma^2}\right) w_q + \sigma \tilde{w}_p\right] \tag{4.45}$$

and

$$\tilde{q} = \frac{1}{D(\sigma)}(\sigma w_q + \tilde{w}_p), \tag{4.46}$$

where

$$D(\sigma) = 1 + \sigma^2 + 2T\sqrt{m^2+\sigma^2} \tag{4.47}$$

and

$$\tilde{w}_p = w_p + T\int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma^2}|x'|}(\sigma w_u(x') + w_v(x'))dx'. \tag{4.48}$$

We also note the following correspondence:

$$\begin{aligned} w_q &= q(t=0) \\ w_p &= p(t=0) \\ w_u &= u(x, t=0) \\ w_v &= v(x, t=0). \end{aligned}$$

Having derived solutions we proceed to invert the transforms by evaluating the Bromwich integral [65]. The contour involved, which runs along either side of each branch cut, may be carefully manipulated so that it encircles spurious poles, and hence makes the contributions from the Landau poles completely apparent. Such a deformation procedure is illustrated in figure 4.7.

We recall that the poles are located where $D(\sigma)$ has zeros, i.e.

$$\sigma^* = \pm\sqrt{-1 + 2T^2 \pm 2T\sqrt{m^2 - 1 + T^2}}. \tag{4.49}$$

For short times the contribution of these Landau poles will dominate the growth rate, which is actually a rate of decay. We expect a decay rate of the form

$$p(t) \propto e^{-{\rm Re}\sigma^* t}, \tag{4.50}$$

and hence from (4.44) we estimate that the energy will decay like $\exp(-2{\rm Re}\sigma^* t)$.

Consider $q(t)$ first, and for simplicity let the string and oscillator be initially in the rest position, i.e. $u(x, t=0) = w_u = 0$, $q(t=0) = w_q = 0$. Then from (4.46) and (4.48),

$$q(t) = \oint \frac{e^{\sigma t}}{D(\sigma)}\left[w_p + T\int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma^2}|x'|} w_v(x')dx'\right] d\sigma. \tag{4.51}$$

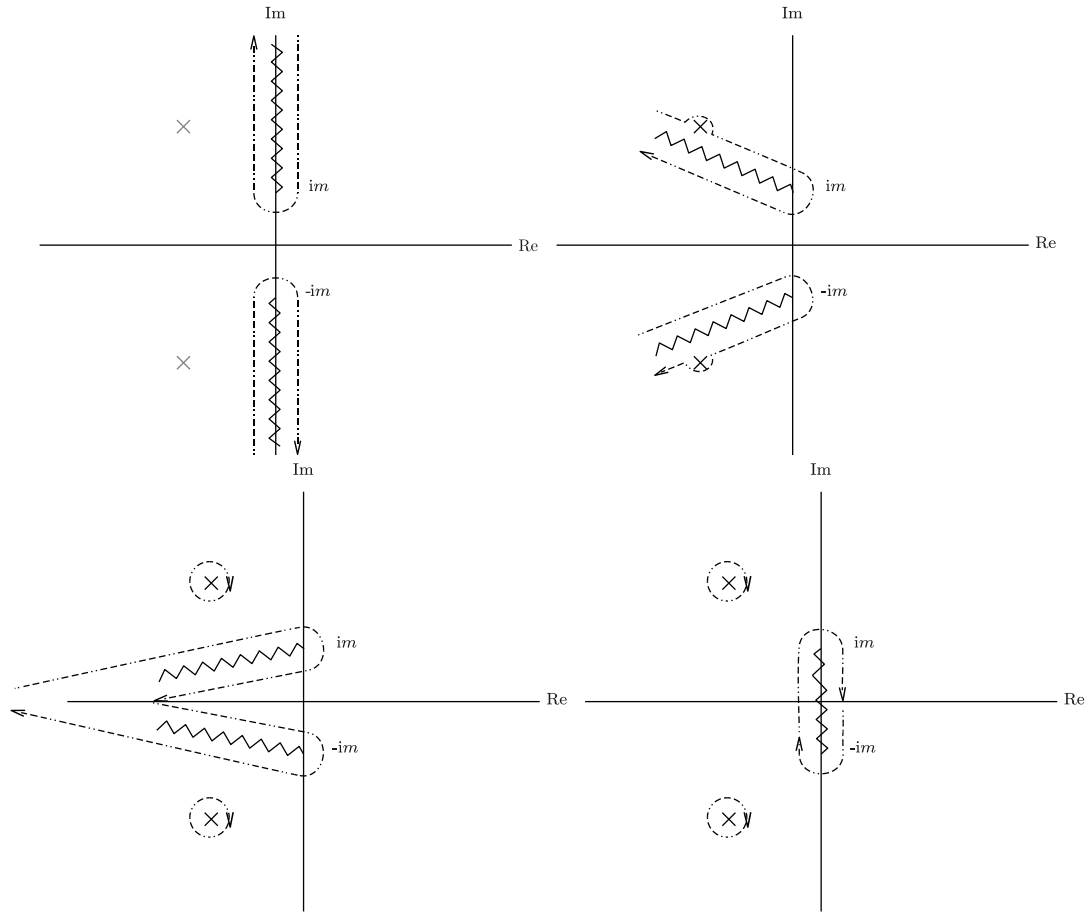


Figure 4.7: We begin with the original integration path over a large circle in the complex plane, deformed such that the contour runs along either side of the branch cut (top left panel). The grey x shows the position of a Landau pole, which is not thus far "seen" by the integration contour, as it lies on the "wrong" Riemann sheet. Next, the contours are rotated about the branch points. By continuing this process we eventually uncover two poles (top right panel), that "switch on" as the integration contour passes over them — indicated by their being marked in black. Next the integration contours are joined up, and two individual contours are left as remnants around the poles. Finally in the bottom right panel, we bring the deformation to its natural conclusion by merging the cuts and contours to lie once more along the imaginary axis.

The integral comprises two contours — one on either side of the branch cut, parametrised using the variable s — and most importantly, residues from the Landau poles. Hence the result is

$$q(t) = \int_{-m}^{m} \tilde{q}(s)\mathrm{e}^{\mathrm{i}s}ds - \int_{-m}^{m} \tilde{q}(s)\mathrm{e}^{-\mathrm{i}s}ds + 2\pi \sum \mathrm{e}^{\mathrm{i}\sigma^*}, \tag{4.52}$$

and so we expect

$$\mathcal{O}(q(t)) = \mathcal{O}(t^{-m}) + \mathcal{O}(\mathrm{e}^{-t}). \tag{4.53}$$

As a result, the decay of energy over time should have two distinct phases. Initially, the exponential decay due to Landau damping will be evident, but at some point the more gentle decay due to the continuous spectrum will begin to dominate. The size and contribution of the branch cut is dependent on m, and so for the Landau damping to be measurable it is necessary to take $m \ll 1$. Qualitatively, increasing m results in the changeover between exponential and algebraic decay starting earlier in the solution history.

Proceeding with the calculation, we integrate along either side of the branch, using the parametrisation $\sigma = is$, and $x = iy$. With this choice we have

$$D(\sigma) = 1 - s^2 + 2T\sqrt{m^2 - s^2}. \tag{4.54}$$

Also, $\bar{D}(\sigma)$ is the function on the other side of the branch cut, i.e.

$$\bar{D(\sigma)} = 1 - s^2 - 2T\sqrt{m^2 - s^2}. \tag{4.55}$$

Now from 4.47:

$$\begin{aligned} q(t) &= \int_{-m}^{m} e^{\mathrm{i}st} w_p \left(\frac{1}{D(\mathrm{i}s)} - \frac{1}{\bar{D}(\mathrm{i}s)} \right) ds \\ &= \int_{-m}^{m} e^{\mathrm{i}st} w_p \left[\frac{4T\sqrt{m^2 - s^2}}{1 - 2s^2 + s^4 - 4T^2(m^2 - s^2)} \right] ds. \end{aligned}$$

Let us consider $m \ll 1$, $t = \mathcal{O}(m^{-1})$, and define

$$\tau = tm, \quad S = s/m. \tag{4.56}$$

From what has gone before we construct

$$I_{q1} = m \int_{-1}^{1} e^{\mathrm{i}S\tau} w_p \left[\frac{4Tm\sqrt{1 - S^2}}{1 - 2S^2m^2 + S^4m^4 - 4T^2m^2(S^2 - 1)} \right] dS. \tag{4.57}$$

The main contribution is

$$\begin{aligned} I_{q2} &= m \int_{-1}^{1} e^{iS\tau} w_p (4Tm\sqrt{1 - S^2}) dS \\ &= \frac{4Tm^2 w_p \pi J_1(\tau)}{\tau} \\ &= \frac{4Tm w_p \pi J_1(mt)}{t}. \end{aligned}$$

For $t \gg 1$,

$$J_1(t) \approx \sqrt{\frac{2}{\pi t}} \cos\left(t - \frac{3\pi}{4}\right); \tag{4.58}$$

hence for long times,

$$I_{q2} \propto t^{-3/2}. \tag{4.59}$$

Using similar reasoning, we can arrive at an approximate decay rate for $p(t)$ in the long time limit:

$$\begin{aligned} I_{p2} &= \int_{-1}^{1} \mathrm{i}Se^{\mathrm{i}S\tau} w_p(4Tm\sqrt{1-S^2})ds \\ &= \frac{-4Tmw_p\pi(\tau J_0(\tau) - 2J_1(\tau))}{\tau^2} \\ &= -4Tw_p\pi\left[\frac{J_0(t)}{t} - \frac{2J_1(mt)}{mt^2}\right]. \end{aligned}$$

For long times, the first term within the square brackets dominates, and since for $t \gg 1$

$$J_0(t) \approx \sqrt{\frac{2}{\pi t}} \cos\left(t - \frac{\pi}{4}\right), \tag{4.60}$$

then for long times

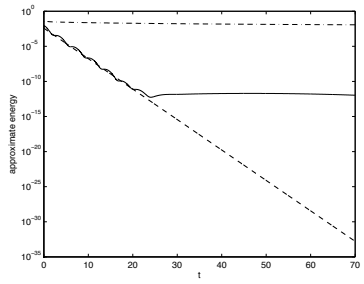
$$I_{p2} \propto t^{-3/2}. \tag{4.61}$$

Since both $p(t)$ and $q(t)$ eventually decay like $t^{-3/2}$, then we expect $E_{\mathrm{osc}} \propto t^{-3}$ after the time at which the contribution from the Landau poles ceases to be the dominant one.

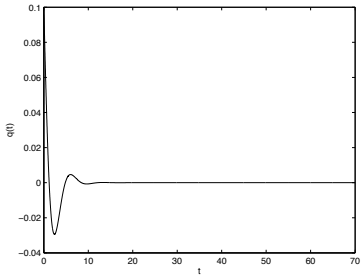
4.3.5 Numerical experiments

In this section, the growth rates estimated in §4.3.4 are tested against numerically computed solutions of the model equations. In all the following numerical experiments, the oscillator and string are started from rest, and the oscillator excited with an instantaneous "kick". The oscillator energy can then be approximated by $E_{\mathrm{osc}} = p^2/2 + q^2/2$. We consider the regimes of m and T as described in section E.1.

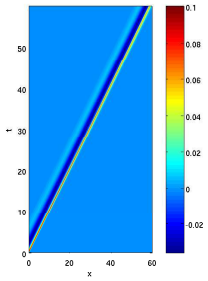
The numerical scheme is as follows. A simple centred second-order in space, leapfrog in time finite-difference scheme was implemented, together with an Asselin filter to damp the computational mode inherent in leapfrog schemes. The filter damps the high frequency computational mode very successfully, whilst degrading the total energy very little. The effect of the filter was checked by computing the total energy (4.43) using a trapezium rule integration scheme. Wave radiation to infinity on the string was handled using the exact DNRBC (without truncation) from §2.3.3 of chapter 2, with coefficients calculated by the DFT method. As with numerical experiments in chapter 3, the runs are short enough to make storing the entire solution history (necessary for the DNRBC) practical. In all experiments $\Delta x = 0.02$ and $\Delta t = 0.01$.



(a) $E_{\mathrm{osc}(t)}$ (solid line), theoretical prediction (dashed line) and total energy (dot-dashed line).



(b) $q(t)$.



(c) $u(x,t)$.

Figure 4.8: Results for $m = 0.05$, $T = 0.5$ (region 1), with $\Delta x = 0.02$ and $\Delta t = 0.01$.

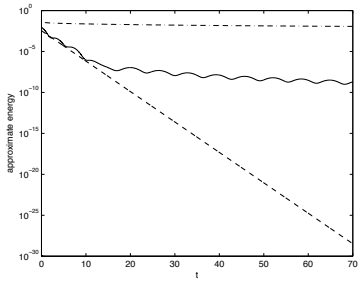
Region 1

Recall that in this region there are four Landau poles, all with non-zero real and imaginary parts. Figure 4.8a shows the approximated oscillator energy from the simulation, together with the theoretical decay rate due to the Landau pole. Here $m = 0.05$ and $T = 0.5$: equation (4.49) for the Landau poles gives a pole with a real part of 0.49937, and hence the predicted decay rate for early times (before algebraic decay takes over) is $\exp(-0.9987t)$. As shown by the plot, the agreement is very good. The upper curve also shows that the addition of the filter with filtering constant $\gamma = 0.02$ has a negligible impact on the total energy. We see a dramatic change in behaviour at around $t = 25$, when the continuous spectrum begins to dominate the solution. Figures 4.8b and 4.8c show that the solution motion is heavily damped due to the presence of the Landau pole.

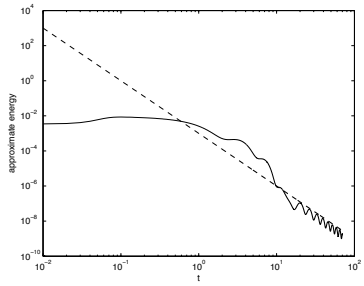
Figure 4.9 shows the results obtained with $m = 0.5$ and $T = 0.5$. The contribution from the continuous spectrum takes over at an earlier time than in the previous experiment. The estimated damping is of the form $E_0 \exp(-0.8856t)$, which is again in good agreement with the calculated energy. In figure 4.9b, the energy has been plotted on a log-log scale, and the curve has the correct t^{-3} decay as predicted from the continuous spectrum.

Region 2

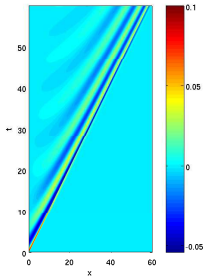
If we choose $m = 0.5$ and $T = 0.9$ corresponding to region 2, there are four Landau poles that lie on the real axis. Figures 4.10a and 4.10b illustrate the resulting oscillator energy, and figure 4.10c $q(t)$. We again see the early decay due to the Landau poles, followed by power law decay. Both decay rates match the theoretical prediction.



(a) E_{osc} (solid line), Landau damping rate predicted by theory (dashed line), and total energy (dot-dashed line).

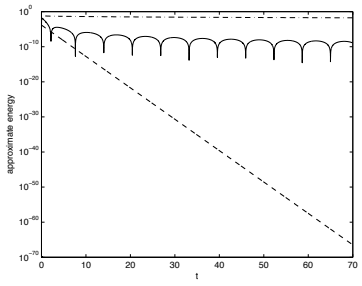


(b) Log-log comparison of E_{osc} and the $E_0 t^{-3}$ damping rate predicted from the continuous spectrum (dashed).

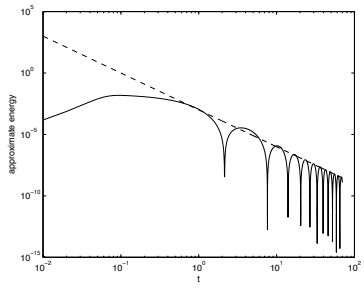


(c) $u(x,t)$.

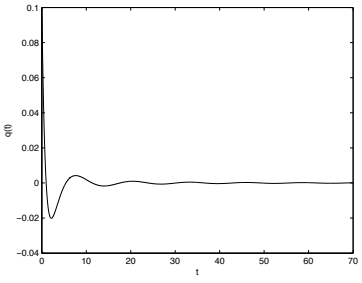
Figure 4.9: Results for $m = 0.5$, $T = 0.5$ (region 1).



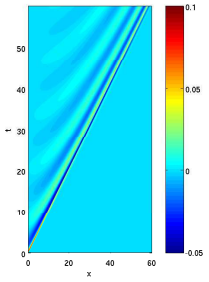
(a) E_{osc} (solid line) and theoretical (dashed) Landau damping rate, and total energy (dot-dashed line).



(b) Log-log plot of E_{osc} (solid line) and algebraic decay due to the continuous spectrum.

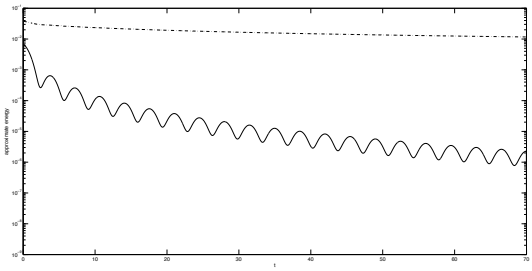


(c) $q(t)$.

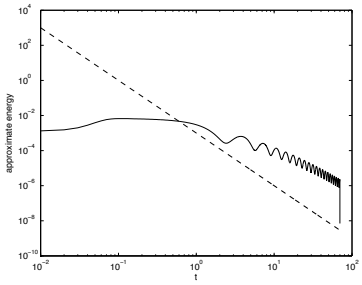


(d) $u(x,t)$.

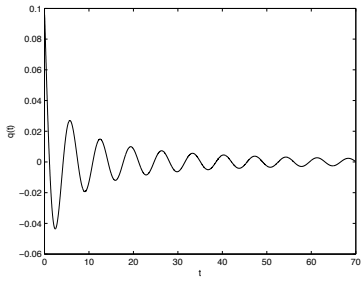
Figure 4.10: Results for $m = 0.5$, $T = 0.9$ (region 2).



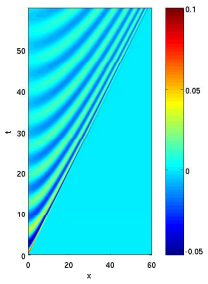
(a) E_{osc} (solid curve) and total system energy (dot-dashed curve).



(b) Log-log plot of E_{osc} (solid curve) and the estimated $E_0 t^{-3}$ decay rate; $m = \mathcal{O}(1)$ gives rise to the disagreement.



(c) $q(t)$. The reduction in damping is clear.



(d) $u(x,t)$.

Figure 4.11: Results for $m = 0.9$, $T = 0.5$ (region 3).

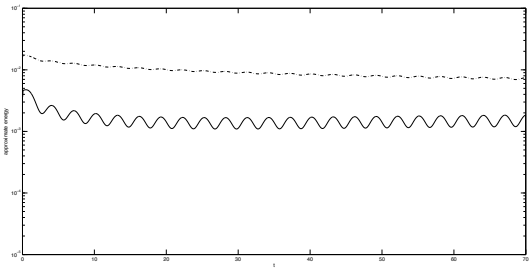
Region 3

We take $m = 0.9$ and $T = 0.5$. Since $m^2 < 1$ we expect all the poles to be Landau poles. All four poles lie on the imaginary axis, and consequently induce no decay in our solution, only undamped oscillation. Furthermore, since the long-time behaviour like t^{-3} is derived in the small m limit, it is reasonable there will be some break-down between theory and agreement, since here m is far from small. Figure 4.11a shows the absence of any exponential decay in the solution energy. Figure 4.11b shows the solution energy on a log-log scale, and we find that it is no longer accurately described by a t^{-3} power law. Figures 4.11c shows the oscillator position; the lack of early damping is evident in the amplitude of the oscillation for large times. The string's displacement, $u(x,t)$, is plotted in the x-t plane in figure 4.11d.

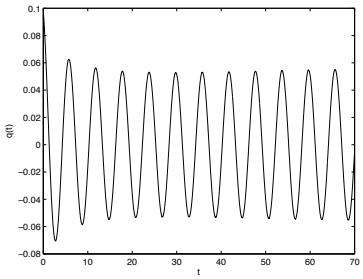
We now take $m > 1$, and consider the region where there are two eigenvalues on the imaginary axis, and two Landau poles. Specifically, we use $m = 1.1$ and $T = 0.2$. Figures 4.12 shows that the results exhibit the non-decaying oscillation due to the eigenvalues.

Region 4

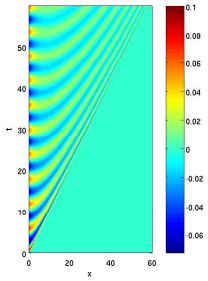
Finally, we consider results for region 4 of the m-T parameter space. Firstly, we look at $m^2 < 1$, taking for example $m = 0.6$, and $T = 0.9$. There are two Landau poles and two eigenvalues: the Landau poles lie on the real axis and the eigenvalues on the imaginary axis. Figures 4.13 shows results similar to those expected for choices of m and T that lie in region one: the early-time decay is exponential and governed by the real part of the Landau pole on the real axis, whereas for long times the power law behaviour due to the continuous spectrum is seen.



(a) E_{osc} (solid) and total energy (dot-dashed).

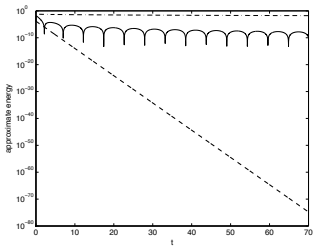


(b) $q(t)$, showing the undamped oscillation clearly.

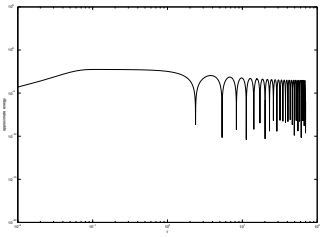


(c) $u(x,t)$.

Figure 4.12: Results for $m = 1.1$, $T = 0.2$.



(a) E_{osc} (solid), predicted decay rate (dashed line) and total energy (dot-dashed).



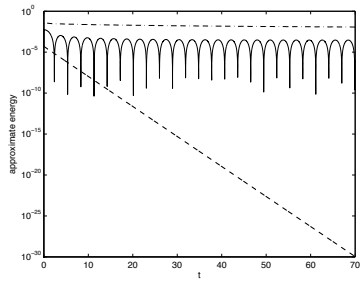
(b) Log-log plot of E_{osc} (solid line), together with $E_0 t^{-3}$ (dashed line).

Figure 4.13: Results for $m = 0.6$, $T = 0.9$ (region 4).

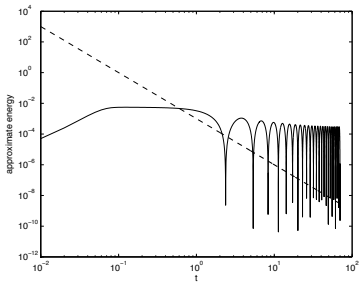
Finally we consider $m^2 > 1$ and the appearance of eigenvalues. We choose $m = 1.1$ and $T = 0.5$. In this region we find two pure real Landau poles, and two pure imaginary eigenvalues. Figure 4.14 shows a solution consistent with the theory; there is a small initial decay, before the dominant constant amplitude oscillation due to the eigenvalue takes over.

4.4 The extended model with damping

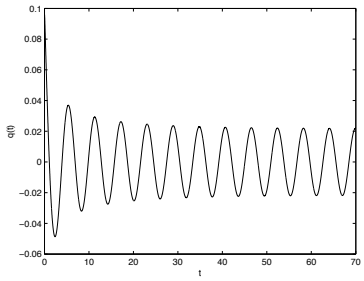
We now consider adding a damping term to the equation governing the string. We choose a damping of the form $+nu_t$ on the left hand side of line two of (4.35), and in doing so transform



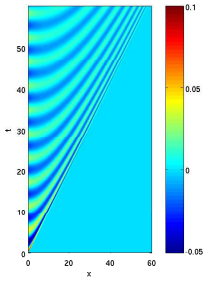
(a) Log plot of $E_{\rm osc}$ (solid line), $E_0 \exp(-\sigma_L t)$ (dashed) and total system energy (dot-dashed line).



(b) Log-log plot of $E_{\rm osc}$.



(c) $q(t)$; the vibrational mode is evident.



(d) $u(x,t)$

Figure 4.14: Results for $m = 1.1$, $T = 0.5$ (region 4).

the KG equation into the telegraph equation. The governing equations for this system are now:

$$\begin{aligned} \ddot{q} + q &= T[u_x]^+_-, \\ u_{tt} - u_{xx} + m^2 u + n u_t &= 0, \\ q(t) &= u(0,t). \end{aligned} \tag{4.62}$$

We again write (4.62) as

$$z_t = \mathcal{L} z,$$

where $z = (p, q, v, u)$, $p = \dot{q}$ and $v = u_t$. Then

$$\mathcal{L} z = \begin{pmatrix} -q + T[u_x]^+_- \\ p \\ u_{xx} - m^2 u - n v \\ v \end{pmatrix}.$$

We obtain the solutions in terms of Laplace transforms as

$$p = \frac{1}{1 + \sigma^2 + 2T\sqrt{m^2 + \sigma(\sigma + n)}} \left[\left(-1 - 2T\sqrt{m^2 + \sigma(\sigma + n)} \right) w_q + \sigma \tilde{w}_p \right] \tag{4.63}$$

and

$$q = \frac{1}{1 + \sigma^2 + 2T\sqrt{m^2 + \sigma(\sigma + n)}} (\sigma w_q + \tilde{w}_p), \tag{4.64}$$

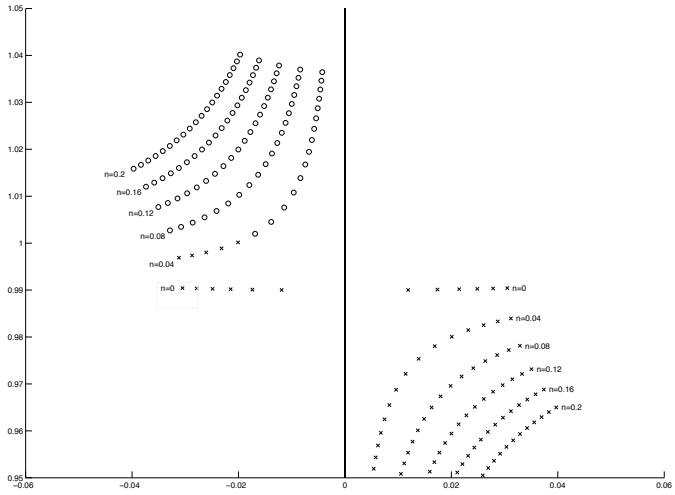


Figure 4.15: Position in the complex plane of the poles (circles) and Landau poles (crosses). These are computed for $T = 0.1$, with n values as shown and $m \in [0.94, 1]$.

where

$$\tilde{w}_p = w_p + T \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma(\sigma+n)}|x'|}[(\sigma+n)w_u + w_v]dx'. \tag{4.65}$$

The values of σ for which the resolvent is not analytic are the solutions of

$$\sigma^4 + (2 - 4T^2)\sigma^2 - 4T^2 n\sigma - 4m^2T^2 + 1 = 0, \tag{4.66}$$

which obviously reduces to (4.49) for $m = 0$. Spurious solutions (connected with Landau poles) correspond to $\mathrm{Re}\left\{\sqrt{m^2 + \sigma(\sigma+n)}\right\} < 0$. Variation of the position and type of the poles as m and n are changed is illustrated by figure 4.15.

As with the undamped case, we may investigate — this time numerically — the positions and nature of each of the poles of $D(\sigma)$ in the m-T plane. The results are shown by the diagram in figure 4.16. The boundary that separates the regions containing "true" poles now curves off to leave a thin region where there are two eigenvalues, for small T. Also we notice that a new region appears where all of the Landau poles are pure real. Finally, the curve that begins at $T = 1, m = 0$, for $n = 0$ has its starting position shifted downwards to a point $T < 1$, by an amount governed by n; in fact this distance is one of the few pieces of information easily obtained analytically, and the relation between n and T at this end point can be derived from the quartic discriminant as

$$n = -\frac{2\sqrt{3}}{9T^2}\sqrt{2T^6 - 3T^4 - 3T^2 + 2 + 2\sqrt{T^12 - 3T^10 + 6T^8 - 7T^6 + 6T^4 - 3T^2 + 1}} \tag{4.67}$$

The bifurcation associated with n becoming nonzero is illustrated in figure 4.17.

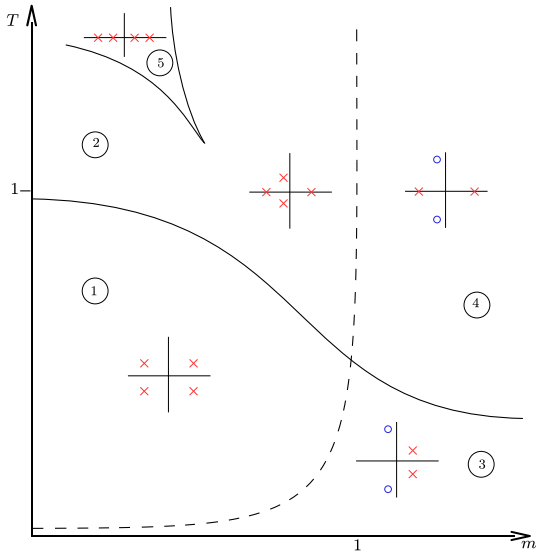


Figure 4.16: Diagram showing the positions in the complex plane of the poles of the resolvent for $n \neq 0$, in each of the four regions in the m-T parameter space. As before, Landau poles are shown as crosses, and eigenvalues as circles.

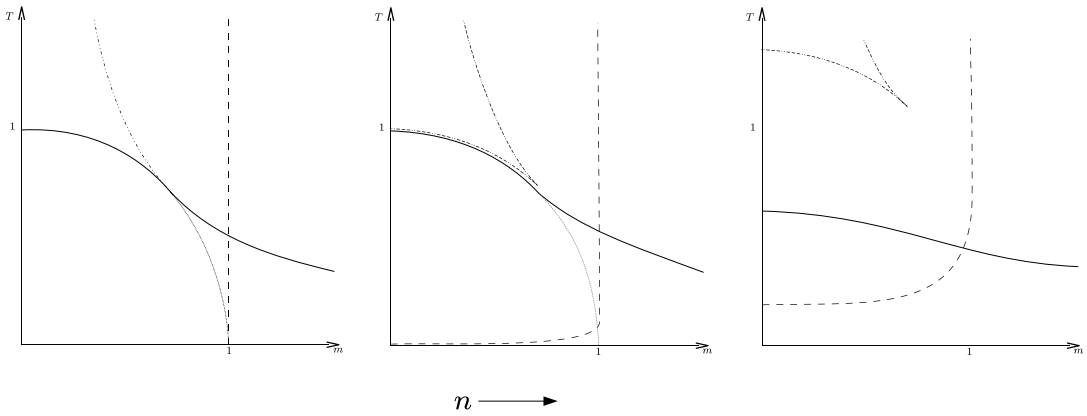


Figure 4.17: Bifurcation diagram associated with nonzero n. Each distinct curve is shown using a different line style.

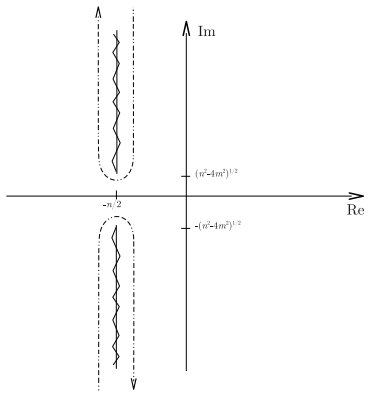


Figure 4.18: Position in the complex plane of the branch points and the chosen branch cuts, for the case of $n < 2m$. The integration contour is also represented schematically by the dashed line.

For $n < 2m$ the branch points have a nonzero imaginary part, and we make cuts on $[-\frac{n}{2} + \sqrt{n^2 - 4m^2}, \infty)$ and $(-\infty, -\frac{n}{2} - \sqrt{n^2 - 4m^2}]$ (fig. 4.18). When integrating we parametrise the contour using s as before

$$\sigma = -n/2 + \mathrm{i}s; \quad s \in [\tfrac{1}{2}\sqrt{4m^2 - n^2}, \infty), \tag{4.68}$$

which changes the form of the dispersion relation

$$\begin{aligned} D(s) &= 1 + \sigma(s)^2 \pm 2T\sqrt{m^2 + \sigma(s)(\sigma(s) + n)} \\ &= 1 + n^2/4 - n\mathrm{i}s - s^2 \pm 2T\mathrm{i}\sqrt{s^2 + n^2/4 - m^2}. \end{aligned}$$

We may now turn to the decay rate in the presence of damping. This proceeds along roughly the same lines as the calculation for $n = 0$. We again attempt to find an estimate for the behaviour of $1/2(p^2 + q^2)$. From before we have that

$$\begin{aligned} q(t) &= \oint \frac{e^{\sigma t}}{D(\sigma)}(\sigma w_q + \tilde{w}_p)\, d\sigma \\ &= \oint \frac{e^{\sigma t}}{D(\sigma)} \left\{ \sigma w_q + w_p + T\int_{-\infty}^{\infty} e^{-\sqrt{m^2 + \sigma(\sigma+n)}|x'|}[(\sigma + n)w_u + w_v]dx' \right\} d\sigma \end{aligned}$$

We now impose $w_u = w_v = w_q = 0$; that is, the string is initially stationary and its displacement zero everywhere, although the oscillator has some initial velocity ($w_p \neq 0$). Then

$$q(t) = \oint \frac{e^{\sigma t}}{D(\sigma)} w_p d\sigma. \tag{4.69}$$

Again the contour of integration goes along each side of the branch, which now lies between the branch points. We parametrise by letting $\sigma = -n/2 + is$, where

$$s \in [-\tfrac{1}{2}\sqrt{4m^2 - n^2}, \tfrac{1}{2}\sqrt{4m^2 - n^2}]. \tag{4.70}$$

Now

$$\begin{aligned} q(t) &= \int_{-\frac{1}{2}\sqrt{4m^2-n^2}}^{\frac{1}{2}\sqrt{4m^2-n^2}} e^{-nt/2+\mathrm{i}st} w_p \left[\frac{1}{D(s)} - \frac{1}{\tilde{D}(s)}\right] ds \\ &= \int_{-\frac{1}{2}\sqrt{4m^2-n^2}}^{\frac{1}{2}\sqrt{4m^2-n^2}} \frac{-4T\mathrm{i}\sqrt{s^2+n^2/4-m^2}e^{-nt/2+\mathrm{i}st}w_p}{1+\frac{n^2}{2}-2s^2-\frac{\mathrm{i}n^2 s}{2}+2\mathrm{i}ns-\frac{3n^2s^2}{2}+4T^2s^2+T^2n^2-4T^2m^2+\frac{n^4}{16}+s^4}. \end{aligned}$$

As before we may let $\tau = tm$ and $S = s/m$. Taking $m \ll 1$, $m/n = \mathcal{O}(1)$ and neglecting small terms we obtain

$$q(t) \approx \frac{-4Tm^2}{1+\frac{n^2}{2}} e^{\frac{-n\tau}{2m}} \int_{-\frac{1}{2}\sqrt{4m^2-n^2}/m}^{\frac{1}{2}\sqrt{4m^2-n^2}/m} e^{\mathrm{i}S\tau} w_q \sqrt{1-n^2/4m^2-S^2}\, dS. \tag{4.71}$$

Similarly, consideration of p yields

$$p(t) \approx \frac{-4Tm^2}{1+\frac{n^2}{2}} e^{\frac{-n\tau}{2m}} \int_{-\frac{1}{2}\sqrt{4m^2-n^2}/m}^{\frac{1}{2}\sqrt{4m^2-n^2}/m} e^{\mathrm{i}S\tau} w_p(-n/2-\mathrm{i}Sm) \sqrt{1-n^2/4m^2-S^2}\, dS. \tag{4.72}$$

Let us Write the integral as

$$I = \int_{-b}^{b} e^{\mathrm{i}st} \sqrt{b^2-s^2} ds, \tag{4.73}$$

where we have renamed S and τ as s and t respectively, and

$$b^2 = 1 - n^2/4m^2. \tag{4.74}$$

From integral tables we find

$$I = \frac{\mathrm{i}bJ_1(bt)}{t}, \quad b > 0. \tag{4.75}$$

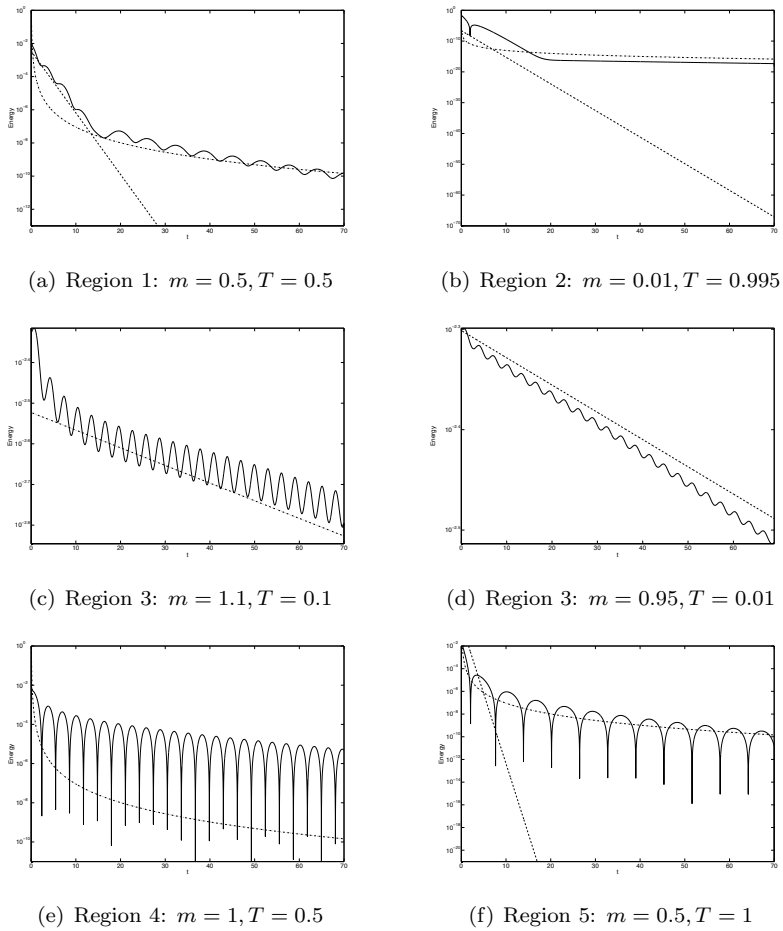
Hence, after substitution and simplification, we find that for long times ($t \gg \frac{3}{4}/\sqrt{1-n^2/m^2}$) the decay rate

$$p(t)^2 \sim \mathrm{e}^{-nt} t^{-3}. \tag{4.76}$$

We have found that the addition of a damping term does not significantly alter the dynamics, in contrast to some other fluid systems where it may change the situation drastically. However, most interesting are the changes in the shapes of the regions of behaviour in the $m - T$ plane; although behaviour in each region is not surprising, a slight change in damping may cause a dramatic change from one region to another. The effects are quite subtle and generally short-lived given the tendency of the damping to bring the system to rest.

4.5 Numerical results: the damped case

We compare the theoretical damping rates with those obtained by numerical experiments. The choice $n = 0.01$ was made to show the influence of damping on the dynamics, whilst not damping the waves too strongly. Other parameters were unchanged from the undamped experiments. Figure 4.19 shows the resulting decay rates and their predictions for regions one to five, as



(a) Region 1: $m = 0.5, T = 0.5$

(b) Region 2: $m = 0.01, T = 0.995$

(c) Region 3: $m = 1.1, T = 0.1$

(d) Region 3: $m = 0.95, T = 0.01$

(e) Region 4: $m = 1, T = 0.5$

(f) Region 5: $m = 0.5, T = 1$

Figure 4.19: Theoretical long-time energy decay rates (dot-dashed curves) and Landau damping rates (dashed curves), together with those found from experiment (solid curves) for each of the five regions in the σ-T parameter space.

illustrated in the diagram in figure 4.16. Agreement is generally very good. Although the damping is relatively modest, the resulting bifurcation in the $m - T$ plane is clear; the second plot for region 3 ($m = 0.95, T = 0.01$) demonstrates the appearance of eigenvalues in the narrow strip across the $m - T$ plane for $n \neq 0$.

4.6 Conclusions

In this chapter we have investigated Landau poles in the context of a simple coupled ODE/PDE system, which has proved a useful prototype for the understanding of Landau damping in more physically realistic (and hence more complicated) settings. The spectral expansion has been derived, together with the solutions in terms of Laplace transforms. From these we have uncovered the eigenvalues and Landau poles of the system, and by deforming the contours of integration have derived growth rate estimates for short timescales, where exponential Landau damping dominates, and for long timescales, where the decay is algebraic. The system was subsequently extended to include a linear damping term. From the analysis we conclude that the damping does not radically alter the dynamics, but does introduce a subtle bifurcation that changes the nature of the solutions in the $m-T$ parameter space, compared to when there is no

damping. The theoretical results agree well with a computed solution, which uses a discretely non-reflecting boundary condition to allow wave radiation to be handled accurately. The results in the chapter provide useful insights into the role of analytic continuation in fluid systems.

Chapter 5

Conclusions and discussion

In this thesis we have studied several aspects of wave radiation relevant to geophysical flows, by means of simplified models that capture some of their features. In parallel, various numerical techniques have been studied and developed, in order to handle the aspects of numerical simulations of wave radiation problems accurately and efficiently.

In chapter 2, non-reflecting boundary conditions were introduced, specifically *discretely* non-reflecting boundary conditions which are derived to provide perfect boundary conditions for a particular discretisation scheme. This chapter paved the way for work in chapters 3 and 4, which both involve schemes where an effective non-reflecting boundary is crucial to the simulations. A DNRBC studied by Han & Zhang was introduced, and several methods for determining its coefficients were given, including a novel method employing the discrete Fourier transform. We examined the limitations of such a boundary condition when applied, which have previously been overlooked: necessary truncation of the infinite coefficient series leads to an inexact condition, which was shown to be suboptimal and even unstable in some cases. It was shown that when using a finite number of coefficients, the optimal boundary condition (in terms of the reflection ratio) is not obtained by truncating the infinite series, but by starting from scratch using an optimisation method. By using GKS theory a simple condition on such coefficients was derived which ensures stability with respect to non-oscillatory modes, and such a condition may be added to the procedure used to derive optimal coefficients. Stability was also examined in more general cases using a Nyquist-type argument. The presence of an Asselin filter was shown to degrade the performance of DNRBCs by rendering them imperfect, and an improved set of coefficients was derived to counteract this. Such an improvement may find application in weather prediction schemes, where accurate NRBCs are required, whilst at the same time the use of the Asselin filter is commonplace.

While only one particular model equation was studied with one particular NRBC, it is possible that some of this work could form a prototype for investigations using other model equations which may be closer to physical reality, with perhaps more advanced schemes, especially where the range of possibilities of methods choosing free parameters in an NRBC is vast. The boundary condition studied provides a good example of practical, seemingly trivial alterations of a scheme (i.e. truncation) having potentially drastic effects on accuracy and even stability, and highlights the need for care in such cases.

The study of non-reflecting boundaries was continued by looking at Higdon higher-order boundary conditions applied to the 2-d rotating shallow water equations. The discrete formulation of such boundary conditions is difficult to derive by hand, so a novel, efficient, easily programmed algorithm was proposed for this purpose. Numerical experiments showed that the boundary condition performs well when implemented using a particular staggered grid scheme, and that it continues to improve as the order of the boundary condition is increased, as expected from theory. The boundary condition features some parameters c_j to be chosen, though setting $c_j = 1 \ \forall j$ was shown to give acceptable results. A comparison was made with a recent method of choosing the coefficients proposed by Givoli & Neta (which has not previously been combined with the Higdon boundary condition applied to this interior numerical scheme), though this was actually shown to perform *less* well. Given that the scheme's stability was not determined previously (as noted by Dea), the GKS-stability of the scheme was examined by a spectra/pseudospectral techniques, by first writing each time step update as a matrix multiplication, determining the general form of the multiplying matrix $\mathbf{A}$, and studying the form of $||\mathbf{A}||^k$ — a technique that has received little attention in the literature. Pseudospectra were studied using standard direct methods and a boundary tracing method due to Kostin and Br uhl. Investigation showed that the absolute stability of the scheme depends on both the grid size and the Coriolis parameter. Even for choices that yield an absolutely stable scheme, the pseudospectrum reveals the potential for milder, transient instability. Despite the drawbacks of such a numerical approach to determining stability, it does have the advantage of providing results almost automatically, with only a reformulation of the problem required. As Trefethen notes [74], it will be interesting to to see if this approach is used more often in the future.

In Chapter 3, we considered wave radiation as a restoring mechanism by means of a simple mixed ODE/PDE model — an extension to the Lorenz-Krishnamurthy system — featuring dispersive wave radiation to infinity, together with some smooth stochastic forcing. Such a model is novel because the dissipation is provided by the radiation rather than a damping mechanism. Balance relations were derived in order to separate inertia-gravity-type waves from the underlying slow dynamics. The system was conjectured to exhibit wave radiation exponentially small in an effective Rossby number, with a wave flux dependent on the parameter controlling the frequency of the forcing. This was confirmed by numerical experiments. Even when the form of the forcing is unknown, we showed it possible to predict the time averaged flux through knowledge of the long-time behaviour of the slow variables, since we expect their power spectra to have an exponential form. In this case, the energy in the power spectrum of the pendulum angle at the frequency $1/\epsilon$ was shown to correlate well with the flux. Further, the balance relations make it possible to diagnose the energy of the fast waves. The numerical scheme used features a non-reflecting boundary implemented using the exact DNRBC from chapter 2, computed using the DFT method and applied without truncation. Results from this model go some way to explaining how wave radiation can act as a dissipative mechanism that ensures a fluid system remains balanced, and perhaps suggest that I-G wave radiation is not a significant mechanism behind downscale energy transfer from mesoscale eddies in the ocean.

In Chapter 4, the role of quasimodes in fluid systems was illustrated by means of another simple model — a novel extension of one originally due to Lamb. The resolvent for the system

was derived, and solutions found in terms of Laplace transforms. Landau poles were found by considering roots of the dispersion relation in these solutions, and the placement of the roots examined as the various parameters were varied, leading to the identification of five distinct regions of behaviour in the parameter space. Estimates for the growth rates of these solutions were derived using the asymptotic forms of the relevant solutions. Growth rates obtained numerically (involving numerical schemes requiring the exact DNRBC of chapter 2) in each region of the parameter space were shown to give good agreement with theory, and demonstrated the relevant different types of mode. A further alteration to the model was the addition of a Rayleigh damping term, which, while naturally damping the solutions, also introduces a subtle bifurcation which alters the solution regions in the parameter space. The aim of this chapter was to try and further the understanding of quasimodes and their associated Landau poles in relevant fluid systems, by means of a model that allows us to probe it analytically, rather than having to resort to numerical means as is so often the case due to nonlinearity. Although the mechanism by which Landau damping arises in the model (dissipation by wave radiation) is different to that in a real fluid, such a model provides us with useful insights when dealing with more physically motivated systems for which the calculations are somewhat unwieldy.

Appendix A

The structure of the update matrix A for the Higdon NRBC.

The form of the sparse square matrix $\mathbf{A}$ is determined by the difference scheme together with the particular arrangement of the storage vector. Algorithm 3 shows how the entire update for all three fields can be decomposed into four separate steps, namely a height field evolution step, application of the height NRBC, a velocity evolution step, a final application of the velocity NRBC, and finally a shift in memory. This procedure can be summarised using five equations, and by writing $\boldsymbol{\phi}_J^n = ((\boldsymbol{\eta}_J^n)^T, (\mathbf{u}_J^{n-\frac{1}{2}})^T, (\mathbf{v}_J^{n-\frac{1}{2}})^T)^T$, where a superscript $n, n-J$ indicates that the column vector contains the field evaluated at the current time, and data as far back as J previous time levels, and hats denote a field that has yet to have the boundary condition applied.

$$\mathbf{U}\begin{bmatrix} \boldsymbol{\eta}^{n,n-J} \\ \mathbf{u}^{n-\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n-\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}^{n,n-J} \\ \hat{\mathbf{u}}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \hat{\mathbf{v}}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix}, \tag{A.1a}$$

$$\mathbf{U_b}\begin{bmatrix} \boldsymbol{\eta}^{n,n-J} \\ \hat{\mathbf{u}}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \hat{\mathbf{v}}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}^{n,n-J} \\ \mathbf{u}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix}, \tag{A.1b}$$

$$\mathbf{H}\begin{bmatrix} \boldsymbol{\eta}_J^{n,n-J} \\ \mathbf{u}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\eta}}_J^{n+1,n-J} \\ \mathbf{u}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix}, \tag{A.1c}$$

$$\mathbf{H_b}\begin{bmatrix} \hat{\boldsymbol{\eta}}_J^{n+1,n-J} \\ \mathbf{u}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}_J^{n+1,n-J} \\ \mathbf{u}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix}, \tag{A.1d}$$

$$\mathbf{B}\begin{bmatrix} \boldsymbol{\eta}_J^{n+1,n-J} \\ \mathbf{u}^{n+\frac{1}{2},n-\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n-\frac{1}{2}-J} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}_J^{n+1,n+1-J} \\ \mathbf{u}^{n+\frac{1}{2},n+\frac{1}{2}-J} \\ \mathbf{v}^{n+\frac{1}{2},n+\frac{1}{2}-J} \end{bmatrix}. \tag{A.1e}$$

After determining the forms of the square matrices $\mathbf{U}, \mathbf{U_b}, \mathbf{H}$ and $\mathbf{H_b}$, and $\mathbf{B}$ we simply form the *update matrix*, $\mathbf{A}$:

$$\mathbf{A} = \mathbf{B}\mathbf{H_b}\mathbf{H}\mathbf{U_b}\mathbf{U}. \tag{A.2}$$

Deriving the structures of the above matrices is an involved, though entirely mechanical process. The storage vector was formed by beginning at the top left corner of the staggered grid, and splicing together rows of the grid, so an entry for h is followed by u, then v, then h etc. Then the data from earlier time levels is added beneath. If $M[N]$ is the total number of grid points in the $i[j]$-direction, then $i = 0, \frac{1}{2}, 1, \ldots, i_{\max}$ where $i_{\max} = \lceil M/2 \rceil$, and $j = 0, \frac{1}{2}, 1, \ldots, j_{\max}$, where $j_{\max} = \lceil N/2 \rceil$. For simplicity and consistency we will deal always with grids that have both M and N odd, and with the four corner points occupied by values of η. Thus, the memory vector has the following form:

$$(\overbrace{\eta^n_{0,0}, u^n_{0,\frac{1}{2}}, v^n_{0,\frac{1}{2}}, \eta^n_{0,1}, \ldots, v^n_{i_{\max},j_{\max}-\frac{1}{2}}, \eta^n_{i_{\max},j_{\max}}}, \overbrace{\eta^{n-1}_{0,0}, \ldots, \eta^{n-1}_{i_{\max},j_{\max}}}, \ldots, \overbrace{\eta^{n-J}_{0,0}, \ldots, \eta^{n-J}_{i_{\max},j_{\max}}})^T. \tag{A.3}$$

The braces above are used to group together individual time levels. It is easy to check that there are a total of $3N(M-1)/2 + N + \lfloor N/2 \rfloor$ values recorded at each time level, hence the storage vector has a total of $(J+1)(3N(M-1)/2 + N + \lfloor N/2 \rfloor)$ entries.

The matrices $\mathbf{H}$ and $\mathbf{U}$ have the following structure

$$\left[\begin{array}{ccccccccccccccccc}
1 & & & & & & & & & & & & & & & & \\
& \ddots & & & & & & & & & & & & & & & \\
& & 1 & & & & & & & & & & & & & & \\
\hline
& & & \mathbf{X} & & & & & & & & & & & & & \\
& & & & 1 & & & & & & & & & & & & \\
& & & & & 1 & & & & & & & & & & & \\
& & & & & & 1 & & & & & & & & & & \\
& & & & & & & \mathbf{Y} & & & & & & & & & \\
& & & & & & & & 1 & & & & & & & & \\
& & & & & & & & & 1 & & & & & & & \\
& & & & & & & & & & 1 & & & & & & \\
& & & & & & & & & & & \mathbf{X} & & & & & \\
& & & & & & & & & & & & \ddots & & & & \\
& & & & & & & & & & & & & \mathbf{Y} & & & \\
\hline
& & & & & & & & & & & & & & 1 & & \\
& & & & & & & & & & & & & & & \ddots & \\
& & & & & & & & & & & & & & & & 1
\end{array}\right].$$

The top block features a total of $N + \lfloor N/2 \rfloor + 2$ ones on the main diagonal, while the bottom block has the same number plus a further $J(3N(M-1)/2 + N + \lfloor N/2 \rfloor)$ ones, also on the main diagonal. The middle block consists of a total of $N-2$ submatrices, separated by groups of three ones. These submatrices, $\mathbf{X}$ and $\mathbf{Y}$, are given in terms of three row vectors, $\mathbf{p}, \mathbf{q}$, and $\mathbf{r}$

as follows

$$\mathbf{X} = (N + \lceil N/2 \rceil - 4) \text{ rows} \left\{ \begin{bmatrix} \mu\mathbf{p} \\ (1-\mu)\mathbf{q} \\ (1-\mu)\mathbf{r} \\ \mu\mathbf{p} \\ (1-\mu)\mathbf{q} \\ \vdots \\ (1-\mu)\mathbf{r} \end{bmatrix} \right. ; \quad \mathbf{Y} = (N + \lfloor N/2 \rfloor - 2) \text{ rows} \left\{ \begin{bmatrix} (1-\mu)\mathbf{q} \\ (1-\mu)\mathbf{r} \\ \mu\mathbf{p} \\ (1-\mu)\mathbf{q} \\ (1-\mu)\mathbf{r} \\ \vdots \\ \mu\mathbf{p} \end{bmatrix} \right. , \quad \text{(A.4)}$$

where

$$\mu = \begin{cases} 1 \text{ for } \mathbf{H} \\ 0 \text{ for } \mathbf{U} \end{cases} \quad \text{(A.5)}$$

and with

$$\mathbf{p} = (0, \ldots, 0, -H\lambda_x, \overbrace{0, \ldots, 0, -H\lambda_y}^{N+\lceil N/2 \rceil - 1}, \mathbf{1}, \overbrace{0, H\lambda_y, 0, \ldots, 0}^{N+\lfloor N/2 \rfloor - 1}, H\lambda_x, 0, \ldots, 0), \quad \text{(A.6)}$$

$$\mathbf{q} = (0, \ldots, 0, -\lambda_x, \overbrace{0, \ldots, 0}^{N+\lfloor N/2 \rfloor - 1}, \mathbf{1}, \overbrace{\Delta t, 0, \ldots, 0}^{N+\lceil N/2 \rceil - 1}, \lambda_x, 0, \ldots, 0), \quad \text{(A.7)}$$

$$\mathbf{r} = (0, \ldots, 0, \lambda_y, -f\Delta t, \mathbf{1}, -\lambda_y, 0, \ldots, 0). \quad \text{(A.8)}$$

Here $\lambda_x = \Delta t/\Delta x$, $\lambda_y = \Delta t/\Delta y$, and the bold face 1 corresponds to the entry on the diagonal in the full matrix (**H** or **U**).

A hand-derivation of the general boundary update matrices $\mathbf{H_b}$ and $\mathbf{U_b}$ will not be pursued here; it is straightforward to find them computationally. The memory update matrix is given by

$$\mathbf{B} = \left[\begin{array}{c|c} \mathbf{I}_T & \mathbf{0}_{(J-1)\times T} \\ \hline \multicolumn{2}{c}{\mathbf{I}_{J\times T}} \end{array} \right], \quad \text{(A.9)}$$

where

$$T = 3N(M-1)/2 + N + \lfloor N/2 \rfloor \quad \text{(A.10)}$$

is the total number of values stored at each time level, and **I** and **0** are respectively the identity matrix and zero square matrix.

The derived update matrix was checked against a code implementing the staggered scheme in a more standard fashion. Although the resulting matrix **A** is very large, it is also sparse; figure A.1 shows the pattern of the nonzero entries of one example with $M = N = 11$, and $J = 2$. This gives a matrix with a total of $294,849$ entries, of which only $4,335$ are nonzero — a sparsity of 0.0147. Hence, sparse linear algebra methods come to the rescue, allowing us to tackle such a large eigenvalue/pseudospectrum problem.

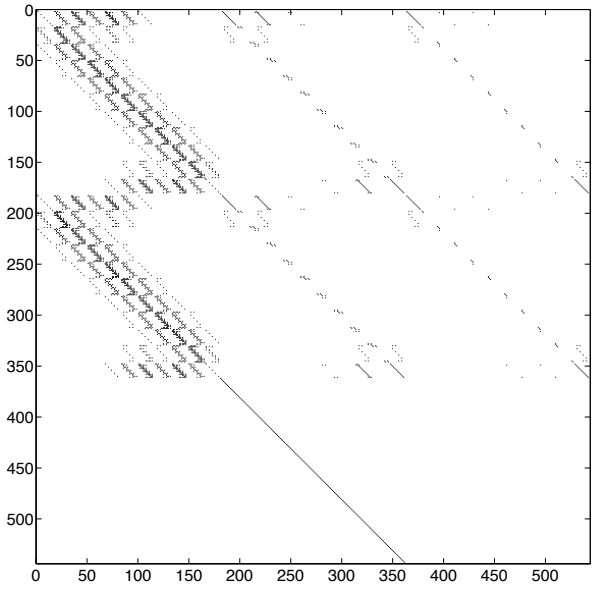


Figure A.1: The sparsity pattern of $\mathbf{A}$, for $M = N = 11$ and $J = 2$.

Appendix B

Basic states and numerical solution of the linearised rotating shallow water equations

In order to study perturbations to an axisymmetric vortex in 2d nonlinear shallow water, we require the steady-state height and velocity fields that correspond to a potential vorticity distribution with a particular profile. In this section we derive the defining equations, and then use numerical techniques outlined by Ford [24] and Schecter et al [69] for determining a basic state given that is the solution to these equations, given a general PV distribution $\bar{q}(r)$. In addition, the linear time evolution problem for disturbances to the basic state is formulated as a system of three PDEs, which may be solved given the computed steady-state fields and initial conditions.

B.1 Model equations

We begin with the rotating shallow water equations:

$$\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + f\hat{\mathbf{z}}\times\mathbf{u} + \nabla\phi = 0, \tag{B.1}$$

$$\partial_t + \nabla\cdot(\phi\mathbf{u}) = 0, \tag{B.2}$$

where ϕ is the geopotential, $\mathbf{u} = (u, v)$ the horizontal velocity vector, and f the Coriolis parameter. The subscript x denotes that the vector components use Cartesian coordinates. We wish to study basic states that are axisymmetric vortices, and hence introduce the radial and azimuthal coordinates r and θ, respectively. Writing the rotating shallow water equations in cylindrical polar coordinates, and defining $u = \mathbf{u}\cdot\hat{\mathbf{r}}$ and $v = \mathbf{u}\cdot\hat{\theta}$, we obtain

$$\partial_t u + (u\partial_r + \frac{v}{r}\partial_\theta)u - \frac{v^2}{r} - fv + \partial_r\phi = 0, \tag{B.3}$$

$$\partial_t v + (u\partial_r + \frac{v}{r}\partial_\theta)v + \frac{uv}{r} + fu + \frac{\partial_\theta\phi}{r} = 0, \tag{B.4}$$

$$\partial_t\phi + u\partial_r h + \frac{hu}{r} + h\partial_r u + \frac{v\partial_\theta v}{r} + \frac{h\partial_\theta v}{r} + \frac{u}{r} + \partial_r u + \frac{\partial_\theta v}{r} = 0. \tag{B.5}$$

B.1.1 An axisymmetric basic state

We seek equations for disturbances to an axisymmetric basic state, i.e. a steady state whose fields are given by $\bar{u} = 0$, $\bar{v} = v(r)$ and $\bar{\phi} = \phi(r)$, which will all be determined by a potential vorticity profile, $\bar{q}(r)$. Using these conditions together with equations (B.3)–(B.5), we find that such a state satisfies the gradient balance equation:

$$\frac{d\bar{\phi}}{dr} = f\bar{v} + \frac{\bar{v}^2}{r}. \tag{B.6}$$

A second equation stems from the definition of potential vorticity:

$$q = \frac{\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u})}{\phi} = \frac{v/r + \partial_r v - (\partial_\theta u)/r}{\phi}, \tag{B.7}$$

which together with our conditions for a basic state yields

$$\frac{d\bar{v}}{dr} + \frac{\bar{v}}{r} = \bar{q}\bar{\phi} - f. \tag{B.8}$$

Choosing a form of $\bar{q}(r)$ and solving (B.6) and (B.8) will give the required fields $\bar{\phi}$ and $\bar{v}$. Details of the numerical solution procedure used with these equations are given in §B.2.

B.1.2 Disturbances to the basic state

We consider disturbances to the basic state of the form $[u', v', \phi'] = [U(r), V(r), \Phi(r)]\mathrm{e}^{\mathrm{i}m\theta - \mathrm{i}\omega t}$, so that m and ω are respectively the azimuthal mode number and the frequency. The $m = 1$ mode represents a pure lateral translation, and is less interesting, whereas $m \geq 2$ modes correspond to deformation of the vortex. This leads to the following equations for the disturbance variables

$$\begin{aligned}
\mathrm{i}\sigma U - \bar{\xi}V + \frac{d\Phi}{dr} &= 0, &\text{(B.9)}\\
\mathrm{i}\sigma V + \bar{\eta}U + \frac{\mathrm{i}m\Phi}{r} &= 0, &\text{(B.10)}\\
\mathrm{i}\sigma\Phi + \frac{1}{r}\frac{d}{dr}(rU\bar{\phi}) + \frac{\mathrm{i}mV\bar{\phi}}{r} &= 0, &\text{(B.11)}
\end{aligned}$$

where we have defined $\sigma = -\mathrm{i}\omega + \mathrm{i}m\bar{\Omega}$. Variables related to the basic state are given by

$$\begin{aligned}
\bar{\Omega}(r) = \frac{\bar{v}}{r}, \quad \bar{\xi}(r) = 2\bar{\Omega} + f,\\
\bar{\eta}(r) = \bar{\zeta} + f, \quad \bar{\zeta} = \frac{1}{r}\frac{d}{dr}(r\bar{v}). \qquad \text{(B.12)}
\end{aligned}$$

We wish to solve the time evolution problem rather than the spectral problem (as was done by Ford in [24]), so assume disturbances of the form $[u', v', \phi'] = [u, v, \phi]\mathrm{e}^{im\theta}$, which gives the following final disturbance equations:

$$\begin{aligned}
\partial_t u &= -\mathrm{i}m\bar{\Omega}u + \bar{\xi}v - \partial_r\phi, &\text{(B.13)}\\
\partial_t v &= -\mathrm{i}m\bar{\Omega}v - \bar{\eta}u - \frac{\mathrm{i}m\phi}{r}, &\text{(B.14)}\\
\partial_t \phi &= -\mathrm{i}m\bar{\Omega}\phi - \frac{1}{r}\partial_r(r\bar{\phi}u) - \frac{\mathrm{i}m\bar{\phi}v}{r}. &\text{(B.15)}
\end{aligned}$$

Note that here and in everything that follows in this section, u, v, and ϕ will refer to the disturbance variables transformed in θ.

B.1.3 Potential vorticity profiles

The potential vorticity profiles studied are as in Schecter & Montgomery [69]. Specifically, the Gaussian cyclone

$$\bar{q} = \frac{b_1}{f r_o^2}(1 + b_2 \mathrm{e}^{-r^2/r_o^2}), \tag{B.16}$$

and the hyperbolic cyclone

$$\bar{q} = \frac{b_1}{f r_o^2}\left\{1 + \frac{b_2}{2}\left[1 - \tanh\left(\frac{r - r_o}{r_o \Delta}\right)\right]\right\}. \tag{B.17}$$

The parameter r_o is used to control the size of the "core" of the vortex. For the hyperbolic cyclone, Δ is used to control the "smoothness" of potential vorticity distribution: as $\Delta \to 0$, $\bar{q}$ more closely approximates a step. Note that $\bar{q} \to b_1/f r_o^2$ as $r \to \infty$ in both cases.

B.2 Numerical techniques

B.2.1 Numerical procedure for determining the basic state

Recall that the basic state is found by solving (B.6) and (B.8) for a particular choice of potential vorticity profile, $\bar{q}$. The equations are solved on a large domain $r \in [0, R]$, and so boundary conditions for $\bar{\phi}(0)$, $\bar{\phi}(R)$, $\bar{v}(0)$ and $\bar{v}(R)$ are required. We solve this boundary value problem by taking $v(0) = 0$, making initial guesses for $\bar{\phi}(0)$ and $\bar{\phi}(R)$, and from these values deriving (by a process explained shortly) a corresponding value for $\bar{v}(0)$. Using these four values a standard shooting technique is employed, whereby the equations are solved forward[backward] from the left[right] boundary, to a matching point at $r = 1$. The quantity $|\bar{\phi}_\mathrm{l} - \bar{\phi}_\mathrm{r}| + |\bar{v}_\mathrm{l} - \bar{v}_\mathrm{r}|$ — where l and r subscripts denote left and right solutions — is used as a measure of the accuracy of the solution, and subsequent improved guesses are sought by its minimisation. The iteration process is halted when the accuracy measure fall below some tolerance — 10^{-10} is chosen here.

Given initial guesses for $\bar{\phi}(0)$ and $\bar{\phi}(r)$ which are chosen freely, the value of $\bar{v}(R)$ is fixed and must therefore be determined. We begin this calculation by decoupling (B.6) and (B.8) to form individual second-order ODEs for $\bar{\phi}$ and $\bar{v}$:

$$\frac{d^2\bar{\phi}}{dr^2} + \frac{1}{r}\frac{d\bar{\phi}}{dr} + f^2 - f\bar{q}\bar{\phi} = \frac{2\bar{v}}{r}\frac{d\bar{v}}{dr}, \tag{B.18}$$

$$\frac{d^2\bar{v}}{dr^2} + \frac{1}{r}\frac{d\bar{v}}{dr} - \frac{\bar{v}}{r^2} - f\bar{q}\bar{\phi} = \bar{\phi}\frac{d\bar{q}}{dr} + \frac{\bar{q}\bar{v}^2}{r}. \tag{B.19}$$

We consider the solution of these equations at the right-hand side of our domain. For $r \gg 1$ the nonlinear right-hand side terms are small, as is $d\bar{q}/dr$, due to the step nature of the profiles under consideration. Neglecting these terms and substituting the large r limit of $\bar{q}$, it is necessary to solve

$$\frac{d^2\bar{\phi}}{dr^2} + \frac{1}{r}\frac{d\bar{\phi}}{dr} + f^2 - \frac{b_1\bar{\phi}}{r_o^2} = 0, \tag{B.20}$$

$$\frac{d^2\bar{v}}{dr^2} + \frac{1}{r}\frac{d\bar{v}}{dr} - \frac{\bar{v}}{r^2} - \frac{b_1\bar{\phi}}{r_o^2} = 0. \tag{B.21}$$

We find

$$\bar{\phi}(r) = \mathrm{K}_0(\sqrt{b_1/r_o}r) + \frac{r_o^2 f}{b_1}, \bar{v}(r) = \mathrm{K}_1(\sqrt{b_1/r_o}r), \tag{B.22}$$

where $\mathrm{K}_n(\cdot)$ is a modified Bessel function of order n. Hence we may relate $\bar{v}(R)$ to the initial guess for $\bar{\phi}(R)$ using the relation

$$\bar{v}(R) = \frac{\mathrm{K}_1(\sqrt{b_1/r_o}R)}{\mathrm{K}_0(\sqrt{b_1/r_0}R)}\left(\bar{\phi}(R) - \frac{r_o^2 f^2}{b_1}\right). \tag{B.23}$$

Figure B.1 shows the fields obtained by this numerical procedure for the Gaussian potential vorticity profile. Figure B.2 shows the various fields obtained by the scheme for the hyperbolic profile as the parameter Δ controlling the "hardness" of the step in potential vorticity is varied.

B.2.2 Numerical solution of the disturbance equations

Ford [24] continued to reduce the system to a set of ODE boundary value problems, which were solved numerically to uncover growth rates of disturbances to the vortex. On other hand, Schecter and Montgomery [69] used numerical techniques to search for Landau poles corresponding to quasimodes of the system, and checked results against full nonlinear simulations. The idea of the current derivation is to reduce the problem to a time evolution problem in 1-d, and proceed to solve in order to determine growth rates, and hence a starting point for the position in the complex plane of the Landau poles.

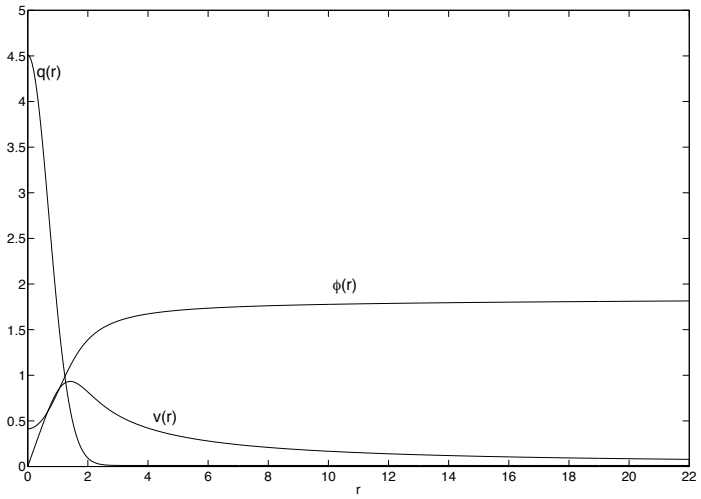
Having obtained the basic states for a particular choice for the form of $\bar{q}(r)$, equations (B.13)–(B.15) are integrated forwards in time using a simple finite-difference scheme, after first adding some initial disturbance to the basic state. An obvious first choice of differencing method would be to use a three point centred-time centred-space stencil, but as we now show this does not yield a useful scheme. Despite the equations involving functions of r multiplying u, v and ϕ, we proceed with the standard stability analysis by first "freezing" these functions at a particular value of r. We follow a stability analysis along the lines of that carried out in the case of the Boussinesq equations, in chapter 3 of [19].

We discretise letting $t = j\Delta t$ and $r = m\Delta r$, writing $u(m\Delta r, j\Delta t) = u_m^j$, and employ three point centred differencing. We examine stability with no basic flow, i.e. $\bar{\Omega} = \bar{\eta} = \bar{\zeta} = 0$ and $\bar{\xi} = f$. The discrete system of equations reads

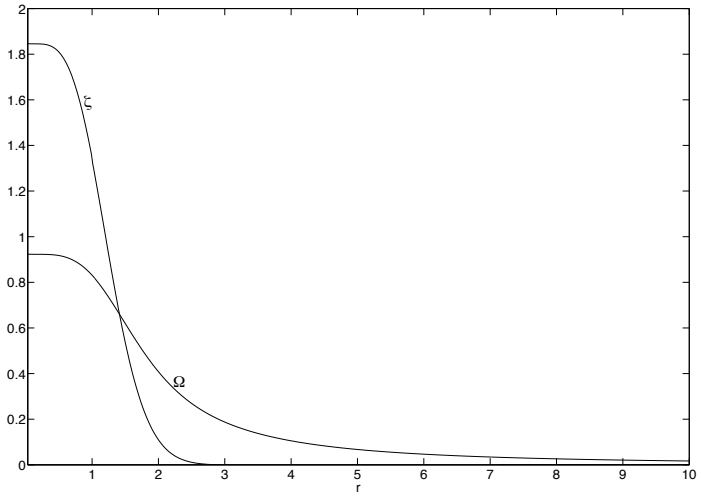
$$\frac{u_m^{j+1} - u_m^{j-1}}{2\Delta t} - f v_m^j + \frac{\phi_{m+1}^j - \phi_{m-1}^j}{2\Delta r} = 0, \tag{B.24}$$

$$\frac{v_m^{j+1} - v_m^{j-1}}{2\Delta t} + \frac{\mathrm{i}m\phi_m^j}{r} = 0, \tag{B.25}$$

$$\frac{\phi_m^{j+1} - \phi_m^{j-1}}{2\Delta t} + \frac{u_m^j}{r} + \frac{u_{m+1}^j - u_{m-1}^j}{2\Delta r} + \frac{\mathrm{i}m v_m^j}{r} = 0. \tag{B.26}$$

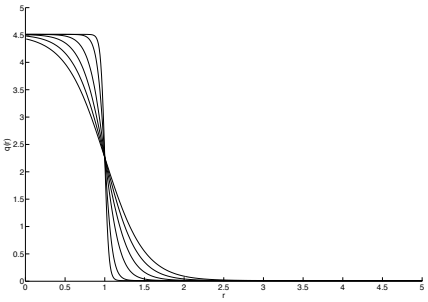


(a) Basic height and azimuthal velocity profiles.

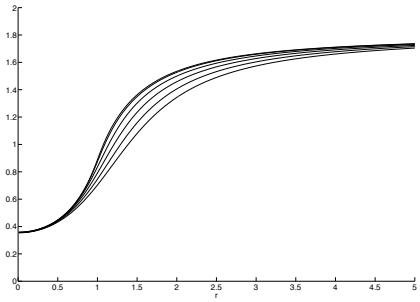


(b) Angular rotation frequency and relative vorticity.

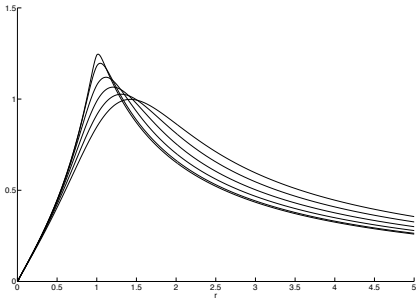
Figure B.1: Fields obtained numerically for the Gaussian PV profile.



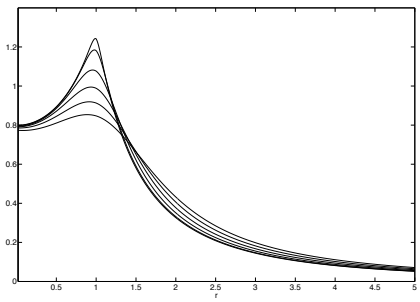
(a) PV profiles. The values chosen for Δ were 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5.



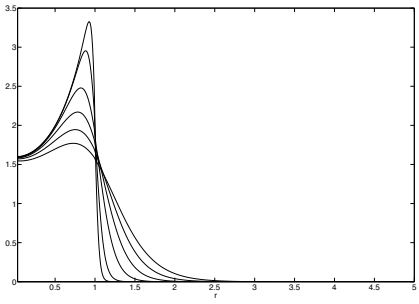
(b) Rest height.



(c) Azimuthal velocity.



(d) Angular rotation rate, Ω.



(e) Relative vorticity, ζ.

Figure B.2: Fields obtained numerically for the hyperbolic PV profile, with $\Delta \in [0.05, 0.5]$.

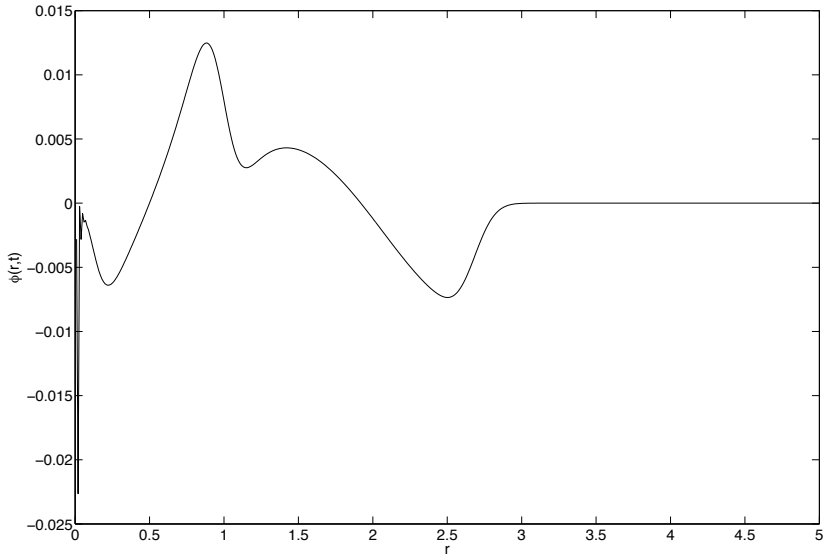


Figure B.3: Appearance of an unstable sawtooth mode in the solution for the height field when using the centred second order stencil.

Substituting solutions of the form $A\mathrm{e}^{\mathrm{i}(k\Delta r-\omega\Delta t)}$, we obtain

$$\begin{bmatrix} -\frac{\mathrm{i}\sin(\omega\Delta t)}{\Delta t} & -f & \frac{\mathrm{i}\sin(k\Delta r)}{\Delta r} \\ 0 & -\frac{\mathrm{i}\sin(\omega\Delta t)}{\Delta t} & \frac{\mathrm{i}m}{r} \\ \frac{\mathrm{i}\sin(k\Delta r)}{\Delta r} & \frac{\mathrm{i}m}{r} & -\frac{\mathrm{i}\sin(\omega\Delta t)}{\Delta t} \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \\ \phi_0 \end{bmatrix} = \mathbf{0}. \tag{B.27}$$

Letting $f = 0$, $r = 1$, $m = 2$ (the lowest mode number of interest), and setting the determinant of the matrix to zero, we find that the discrete dispersion relation obeys

$$\tilde{\omega}(\mathrm{i}\tilde{\omega}^2 + \mathrm{i}\tilde{k}^2 - \tilde{k} - 4\mathrm{i}) = 0. \tag{B.28}$$

where $\tilde{\omega} = \sin(\omega\Delta t)/\Delta t$ and $\tilde{k} = \sin(k\Delta r)/\Delta r$. Hence, $\tilde{\omega}$ clearly has a nonzero imaginary part for real $\tilde{k}$, and so modes will grow exponentially in time. The centred difference scheme is therefore unconditionally unstable in time — a fact confirmed by experiment. Figure B.3 shows an associated sawtooth type mode beginning to make its presence felt in the solution for the height field obtained using the leapfrog scheme; the solution quickly blows up.

We instead employ the third-order Adams-Bashforth time-stepping scheme,

$$y^{j+1} = y^j + \frac{\Delta t}{12}(23F(y^n) - 16F(y^{n-1}) + 5F(y^{n-2})). \tag{B.29}$$

This is found to produce a stable scheme for integrating the disturbance equations for $f = 0$ and any mode number of interest, provided $\lambda = \Delta t/\Delta r$ is sufficiently small.

B.2.3 Boundary conditions

Obviously in this numerical simulation it is necessary that waves generated by perturbations to the basic state are allowed to radiate away to infinity, without being returned. Unfortunately, computations are limited to a finite domain, so some strategy for minimising wave reflection at the boundary must be pursued. On such popular and simple approach is to use a sponge layer or wave absorbing layer [60]. This consists of adding a ring beyond some radius in the computational domain, wherein the velocity fields are subject to linear (i.e. Rayleigh) damping. This corresponds to adding a $-\chi U(r)$ term to the right-hand side of equation (B.13) and $-\chi(r)V$ to the right-hand side of equation (B.14). The continuity equation is undamped, to preserve conservation of mass. The strength of $\chi(r)$ must be chosen such that when it is implemented numerically, the damping is neither too weak to be ineffective, nor to strong such that it effectively acts as a solid wall and returns most of the incident waves. The functional form of $\chi(r)$ can be chosen to improve its absorbing properties. The form we take is quadratic beyond the sponge radius r_{sponge}:

$$\chi(r) = \begin{cases} 0, & 0 < r < r_{\text{sponge}}, \\ \chi_{\max}\left(\frac{r-r_{\text{sponge}}}{R-r_{\text{sponge}}}\right)^2, & r_{\text{sponge}} \leq r \leq R. \end{cases} \tag{B.30}$$

B.2.4 Numerical viscosity

In order to control grid scale artificial noise in the solution, we add some numerical viscosity, i.e. a term of the the form $\nu/r\frac{d}{dr}(rdU/dr) - m^2/r^2$ to the right hand side of (B.13) and a corresponding term to the the equation for V.

B.2.5 Growth rate results

The results produced by the code are confirmed by experiment and comparison with published data. Specifically, a hyperbolic cyclone whose "smoothness" was dependent on a parameter Δ, was perturbed by adding a small initial disturbance in the azimuthal velocity field. The hyperbolic profile is of the form

$$\bar{q} = \frac{b_1}{f r_o}\left\{1 + \frac{b_2}{2}\left[1 - \tanh\left(\frac{r - r_o}{r_o\Delta}\right)\right]\right\}, \quad r < r_b, \tag{B.31}$$

where f is the Coriolis parameter as usual, and r_b is a radius beyond which $\bar{u} = \bar{v} = 0$ and $\bar{\phi} = \text{const.}$, i.e. the vortex is compact. The growth rate of the resulting instability or decay was examined over long time scales. This experiment was carried out using identical parameters and domain size as in §3 of [69], where results are obtained using a spectral code. Specifically, $r_b = 22r_o, b_1 = 2.25 \times 10^{-4}$ and $b_2 = 400$. The Coriolis parameter was "tuned" iteratively so that the resulting Rossby and Froude numbers matched (as closely as could be managed) those quoted, e.g. for $\Delta = 0.1$, $R_{\text{o}} \approx 50$ and $F_{\text{r}} \approx 0.866$. Asymmetry was induced by adding a small perturbation in the velocity fields, with a Gaussian profile centred at the core radius, r_o. Our results are shown in figure B.4. The agreement is good, except for larger values of Δ; this is possibly due to inadequate resolution of fluid activity in the critical layer.

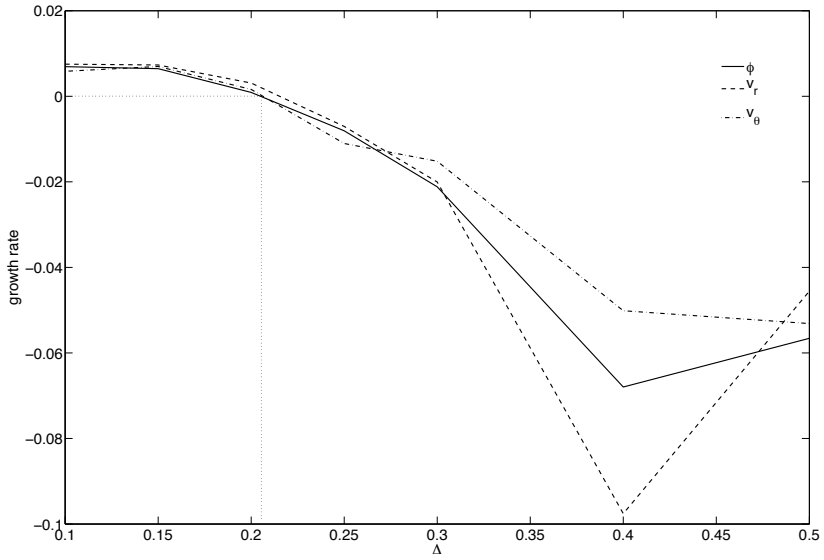


Figure B.4: Growth rates obtained using the 1d linear code, by measuring amplitude of each field at a fixed distance from the centre of the domain.

Appendix C

Numerical solution of the nonlinear rotating shallow water equations

The splitting method developed towards the end of chapter 2 for solving the SWEs was developed to address fairly small scale problems, and so for experiments on a longer timescale and/or higher resolution, a higher performance code is required. The scheme used is closely modelled on the finite-difference scheme described by Sadourny [68], with some minor modifications. This scheme employs an Arakawa "C"-staggered mesh, which has dispersive properties superior to those of an unstaggered mesh, and has the added advantage of precluding the existence of two grid interval noise. The model allows for the addition of a temporally and spatially smooth random surface wind stress. The code features two versions of the evolution equations, allowing the exact numerical conservation of either energy

$$E = \int_\Omega |\mathbf{u}|^2 h/2 + gh^2/2 \; d\Omega \tag{C.1}$$

or enstrophy,

$$Z = \int_\Omega q^2 gh/2d \; \Omega, \tag{C.2}$$

depending on which quantity is most relevant to the problem. The grid is arranged as shown in figure C.1.

Boundary conditions are required for only a single pair of variables on each side of the grid. To write down the differencing scheme, recall the following operators:

$$\Delta_x = \tfrac{1}{d}(\delta_x^{1/2} - \delta_x^{-1/2}), \tag{C.3}$$

where d is the grid spacing (between two points of the same variable) and let the averaging operator (represented by a bar) be defined as

$$\overline{a(x,y)}^{\,x} = \tfrac{1}{2}(\delta_x^{1/2} + \delta_x^{-1/2})a. \tag{C.4}$$

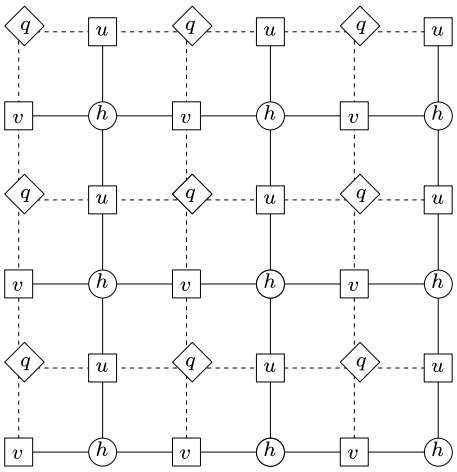


Figure C.1: The staggered grid on which the shallow water equations are solved — this being a very small example.

Then the two schemes may be written in semi-discrete form as follows: first, that which conserves enstrophy

$$\partial_t u - \overline{q\overline{V}^{\,x}}^{\,y} + \Delta_x S = 0, \tag{C.5}$$
$$\partial_t v + \overline{q\overline{U}^{\,y}}^{\,x} + \Delta_y S = 0, \tag{C.6}$$
$$\partial_t h + \Delta_x U + \Delta_y V = 0. \tag{C.7}$$

and second that which conserves energy

$$\partial_t u - \overline{q}^{\,y}\,\overline{\overline{V}^{\,x}}^{\,y} + \Delta_x S = 0, \tag{C.8}$$
$$\partial_t v + \overline{q}^{\,x}\,\overline{\overline{U}^{\,y}}^{\,x} + \Delta_y S = 0, \tag{C.9}$$
$$\partial_t h + \Delta_x U + \Delta_y V = 0. \tag{C.10}$$

In both these sets of equations, the variables U, V, S and q are defined as follows:

$$U = g\overline{h}^{\,x} u, \tag{C.11}$$
$$V = g\overline{h}^{\,y} v, \tag{C.12}$$
$$S = gh + 1/2(\overline{u^2}^{\,x} + \overline{v^2}^{\,y}), \tag{C.13}$$
$$q = \frac{\Delta_x v - \Delta_y u}{g\overline{\overline{h}^{\,x}}^{\,y}}. \tag{C.14}$$

For the time discretisation, a simple choice would be to use a second order leapfrog scheme

$$\phi^n = \phi^{n-2} + 2\Delta t F(\phi^{n-1}). \tag{C.15}$$

However, the computational mode admitted by the leapfrog scheme necessitates time filtering to alleviate sawtooth instabilities with wavelength $2\Delta x$. As a remedy, the code allows the use of the second order Adams-Bashforth scheme — strongly advocated for geophysical problems

by Durran [19]:

$$\phi^n = \phi^{n-1} + \Delta t/2(3F(\phi^{n-1}) - F(\phi^{n-2})). \tag{C.16}$$

For boundary conditions two choices are offered. One may choose to enforce a doubly-periodic boundary condition, or alternatively the approximate yet pragmatic approach taken in so many fluid simulations: *sponge layers*, as described in chapter F. The code implements a discrete flux calculation; since the boundary may be transparent, it is also useful to approximate the total grid energy and enstrophy, using the following

$$E_{\text{tot}} = \tfrac{1}{2}\Sigma g h^2 + g h \overline{u^2}^{\,x} + g h \overline{v^2}^{\,y}, \tag{C.17}$$

$$Z_{\text{tot}} = \tfrac{1}{2}\Sigma g \eta^2 \overline{\overline{h}^{\,x}}^{\,y}. \tag{C.18}$$

The model also calculates the wave flux on a square that lies several grid points within the boundary, using the formula

$$\mathbf{F} = gHh\mathbf{u}. \tag{C.19}$$

The resulting solver is implemented in Parallel using Fortran90 and OpenMP[1], to allow for the use of 16- and 32-core shared memory machines to achieve considerable speedup.

[1] Details of the source files for this solver can be found in the code listing at the end of this thesis.

Appendix D

Higher-order balanced solutions for the eLK model

The second and third order solutions (in terms of Fourier transforms) X_2, Y_2 and X_3 are

$$\begin{aligned} X_2 = \frac{1}{\omega^6}\big[&-12\,f^3w^2uv^3 + 12\,f^3w^2u^3v + 44\,v^3w^2fu - 16\,vw^4fu - 44\,vw^2fu^3 \\ &+uv^5f^3 - 2\,u^3v^3f^3 + 14\,u^3v^3f - uv^5f + u^5vf^3 - u^5vf\big], \end{aligned} \tag{D.1}$$

$$\begin{aligned} Y_2 = \frac{1}{\omega^6}\big[&-f^5v^6w - 20\,f^3v^4w^3 + 3\,f^3v^6w + 44\,v^4w^3f - 16\,v^2w^5f \\ &-v^6wf + u^6f^5w - 20\,u^4f^3w^3 - 3\,u^6f^3w + 16\,u^2w^5f + 44\,u^4w^3f + u^6wf + 3\,f^5v^4wu^2 - \\ &57\,f^3v^4u^2w - 3\,f^5v^2u^4w + 88\,f^3v^2w^3u^2 + 57\,f^3v^2wu^4 + 135\,v^4wu^2f \\ &-135\,v^2wu^4f - 328\,v^2w^3u^2f\big], \end{aligned} \tag{D.2}$$

and

$$\begin{aligned} X_3 = \frac{1}{\omega^8}\big[&-u^7vf^5 + 57\,u^3v^5f^3 - 57\,u^5v^3f^3 - 3\,u^3v^5f^5 + 3\,u^5v^3f^5 \\ &-135\,u^3v^5f + 48\,v^3f^5w^2u^3 + 135\,u^5v^3f - u^7vf - 3\,uv^7f^3 + 3\,u^7vf^3 - 408\,vw^2u^5f \\ &+912\,v^3w^4uf - 408\,v^5w^2fu + 2064\,v^3w^2fu^3 - 288\,v^3f^3w^4u - 840\,v^3f^3w^2u^3 \\ &+228\,v^5f^3w^2u - 24\,vf^5w^2u^5 + 288\,vf^3w^4u^3 + 228\,vf^3w^2u^5 \\ &-64\,vw^6fu - 912\,vw^4fu^3 + uv^7f^5 + uv^7f - 24\,v^5f^5w^2u\big]. \end{aligned} \tag{D.3}$$

The various inverse transforms we require are listed below, up to second order (except y_2, which is omitted due to its length). Solutions x_0 and x_1 may be found in §3.1.3 of chapter 3.

$$\begin{aligned} y_0 = \frac{1}{8\pi}\,A\,wB^2\,(u-v)\,(u+v)\Big(&2\,\sinh(B\,s)\,\mathrm{e}^{1/2\,B^2} \\ &-\mathrm{erf}\left(1/2\,s\sqrt{2} + 1/2\,B\,\sqrt{2}\right)\mathrm{e}^{1/2\,B\,(2\,s+B)} \\ &+\mathrm{erf}\left(-1/2\,s\sqrt{2} + 1/2\,B\,\sqrt{2}\right)\mathrm{e}^{1/2\,B\,(-2\,s+B)}\Big) \end{aligned} \tag{D.4}$$

$$
\begin{aligned}
y_1 = \frac{1}{8\pi} A\, B^3 w \mathrm{e}^{1/2\, B^2} \left(4\, v^2 w^2 + 14\, v^2 u^2 - v^4 - 4\, u^2 w^2 - u^4\right) \\
\times \left(s \cosh\left(B\, s\right) + B\, \sinh\left(B\, s\right)\right) \\
+ \frac{1}{16\pi} A\, B^3 \mathrm{erf}\left(-1/2\, s\sqrt{2} + 1/2\, B\, \sqrt{2}\right) w \mathrm{e}^{1/2\, B\, (-2\, s + B)} \\
\times \left(4\, v^2 w^2 + 14\, v^2 u^2 - v^4 - 4\, u^2 w^2 - u^4\right) \left(-s + B\right) \\
+ \frac{1}{192\pi^3} A^3 B^4 w \sqrt{6}\sqrt{3}\sqrt{2} \left(u - v\right)^2 \left(u + v\right)^2 \\
\times \left(2\, \sinh\left(B\, s\right) + B\, s \cosh\left(B\, s\right) + 3\, B^2 \sinh\left(B\, s\right)\right) \mathrm{e}^{3/2\, B^2} \\
+ \frac{1}{384\pi^3} w A^3 B^4 \sqrt{6}\sqrt{3}\sqrt{2} \mathrm{erf}\left(-1/6\, \sqrt{6} s + 1/2\, \sqrt{6} B\right) \\
\times \left(u - v\right)^2 \left(u + v\right)^2 \left(-B\, s + 3\, B^2 + 2\right) \mathrm{e}^{1/2\, B\, (-2\, s + 3\, B)} \\
+ \frac{1}{16\pi} \frac{A}{} B^3 \mathrm{erf}\left(1/2\, s\sqrt{2} + 1/2\, B\, \sqrt{2}\right) w \\
\times \left(-4\, v^2 w^2 - 14\, v^2 u^2 + v^4 + 4\, u^2 w^2 + u^4\right) \left(B + s\right) \mathrm{e}^{1/2\, B\, (2\, s + B)} \\
- \frac{1}{384\pi^3} w A^3 B^4 \sqrt{6}\sqrt{3}\sqrt{2} \mathrm{erf}\left(1/6\, \sqrt{6} s + 1/2\, \sqrt{6} B\right) \mathrm{e}^{1/2\, B\, (2\, s + 3\, B)} \\
\times \left(u - v\right)^2 \left(u + v\right)^2 \left(3\, B^2 + 2 + B\, s\right)
\end{aligned}
\tag{D.5}
$$

$$
\begin{aligned}
x_2 = \frac{1}{32\pi} b^3 a\, u v \mathrm{e}^{1/2\, b^2} \left(-44\, v^2 w^2 + 16\, w^4 + 44\, u^2 w^2 - 14\, u^2 v^2 + v^4 + u^4\right) \\
\times \left(-b\, s^2 \sinh\left(b\, s\right) - 2\, b^2 s \cosh\left(b\, s\right) - b^3 \sinh\left(b\, s\right) + s \cosh\left(b\, s\right)\right) \\
+ \frac{1}{768\pi^3} u v a^3 b^5 \sqrt{3}\sqrt{2}\sqrt{6} \mathrm{e}^{3/2\, b^2} \left(u - v\right) \left(u + v\right) \left(u^2 + 12\, w^2 - v^2\right) \\
\times \left(12\, b\, \sinh\left(b\, s\right) + 6\, b^2 s \cosh\left(b\, s\right) + 9\, b^3 \sinh\left(b\, s\right) + 3\, s \cosh\left(b\, s\right) + b\, s^2 \sinh\left(b\, s\right)\right) \\
+ \frac{1}{32\pi^{3/2}} b^4 a\, u v s \sqrt{2} \left(-44\, v^2 w^2 + 16\, w^4 + 44\, u^2 w^2 - 14\, u^2 v^2 + v^4 + u^4\right) \mathrm{e}^{-1/2\, s^2} \\
+ \frac{1}{128\pi^{7/2}} u v a^3 b^6 s \sqrt{3}\sqrt{2} \mathrm{e}^{-1/6\, s^2} \left(u - v\right) \left(u + v\right) \left(-u^2 - 12\, w^2 + v^2\right) \\
+ \frac{1}{64\pi} b^3 a\, u v \mathrm{erf}\left(-1/2\, s\sqrt{2} + 1/2\, b\, \sqrt{2}\right) \mathrm{e}^{1/2\, b\, (-2\, s + b)} \\
\times \left(-v^4 + 44\, v^2 w^2 - 44\, u^2 w^2 - u^4 + 14\, u^2 v^2 - 16\, w^4\right) \left(s + s^2 b - 2\, s b^2 + b^3\right) \\
+ \frac{1}{1536\pi^3} \frac{u}{v} a^3 b^5 \sqrt{3}\sqrt{2}\sqrt{6} \mathrm{erf}\left(-1/6\, \sqrt{6} s + 1/2\, \sqrt{6} b\right) \left(u - v\right) \left(u + v\right) \left(u^2 + 12\, w^2 - v^2\right) \\
\times \left(-6\, s b^2 + 9\, b^3 - 3\, s + s^2 b + 12\, b\right) \mathrm{e}^{1/2\, b\, (-2\, s + 3\, b)} \\
+ \frac{1}{64\pi} b^3 a\, u v \mathrm{erf}\left(1/2\, s\sqrt{2} + 1/2\, b\, \sqrt{2}\right) \\
\times \left(-44\, v^2 w^2 + 16\, w^4 + 44\, u^2 w^2 - 14\, u^2 v^2 + v^4 + u^4\right) \\
\times \left(2\, s b^2 + b^3 - s + s^2 b\right) \mathrm{e}^{1/2\, b\, (2\, s + b)} \\
+ \frac{1}{1536\pi^3} u v a^3 b^5 \sqrt{3}\sqrt{2}\sqrt{6} \mathrm{erf}\left(1/6\, \sqrt{6} s + 1/2\, \sqrt{6} b\right) \mathrm{e}^{1/2\, b\, (2\, s + 3\, b)} \\
\times \left(u - v\right) \left(u + v\right) \left(-u^2 - 12\, w^2 + v^2\right) \\
\times \left(6\, s b^2 + 9\, b^3 + 3\, s + s^2 b + 12\, b\right).
\end{aligned}
\tag{D.6}
$$

Appendix E

Calculation of the spectral expansion for the dispersive Lamb model

E.1 Computing the resolvent

We compute the action of the resolvent on an arbitrary vector $w = (w_p, w_q, w_v(x), w_u(x))$. This amounts to solving the system

$$\begin{aligned} \sigma p + q - T\,[u_x]_-^+ &= w_p, && \text{(E.1)} \\ \sigma q - p &= w_q, && \text{(E.2)} \\ \sigma v - u_{xx} + m^2 u &= w_v, && \text{(E.3)} \\ \sigma u - v &= w_u, && \text{(E.4)} \end{aligned}$$

for $z(\sigma) = (p(\sigma), q(\sigma), v(x,\sigma), u(x,\sigma))$. In what follows we omit the dependence on σ. From (E.3)–(E.4), we find that

$$u_{xx} - (m^2 + \sigma^2)u = -(\sigma w_u + w_v).$$

This is best solved by finding the Green's function of the right hand side:

$$G(x,x') = -\frac{1}{2\sqrt{\sigma^2+m^2}} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x-x'|}.$$

As a result the general solution can be written as

$$u(x) = C\mathrm{e}^{-\sqrt{m^2+\sigma^2}|x|} + \frac{1}{2\sqrt{m^2+\sigma^2}} \int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x-x'|} (\sigma w_u(x') + w_v(x'))\, dx', \qquad \text{(E.5)}$$

for some constant C. The choice of the branch of the square root of $m^2 + \sigma^2$ is important. For (E.5) to satisfy radiation boundary conditions as $x \to \pm\infty$ (waves propagate only in the direction from the oscillator to infinity), this square root needs to be such that

$$\mathrm{Re}\,\sqrt{m^2+\sigma^2} > 0. \qquad \text{(E.6)}$$

This dictates the choice of branch cuts $[\mathrm{i}m, \mathrm{i}\infty) \cup (-\mathrm{i}\infty, -\mathrm{i}m]$ for which (E.6) can be verified. Since $u(x=0) = q$, by evaluating (E.5) at $x = 0$ we can show that C is related to q according to

$$C = q - \frac{1}{2\sqrt{m^2+\sigma^2}} \int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x'|} (\sigma w_u(x') + w_v(x'))\, dx' \tag{E.7}$$

Introducing this expression for C into (E.5) gives the solution

$$u = q\mathrm{e}^{-\sqrt{m^2+\sigma^2}|x|} + \frac{1}{2\sqrt{m^2+\sigma^2}} \int_{-\infty}^{\infty} H(x,x';\sigma)(\sigma w_u(x') + w_v(x'))\, dx', \tag{E.8}$$

where for notational convenience we have defined

$$H(x,x';\sigma) = \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x-x'|} - \mathrm{e}^{-\sqrt{m^2+\sigma^2}(|x|+|x'|)}. \tag{E.9}$$

Using (E.8), we compute

$$[u_x]_-^+ = -2\sqrt{m^2+\sigma^2}q + \int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x'|} (\sigma w_u(x') + w_v(x'))\, dx'.$$

Introducing this result into (E.1) gives

$$\sigma p + \left(1 + 2T\sqrt{m^2+\sigma^2}\right) q = \tilde{w}_p, \tag{E.10}$$

where

$$\tilde{w}_p = w_p + T \int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x'|} \left(\sigma w_u(x') + w_v(x')\right)\, dx'.$$

Solving (E.10) and (E.2) for p and q gives

$$\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\sigma^2 + 1 + 2T\sqrt{m^2+\sigma^2}} \begin{pmatrix} \sigma & -1 - 2T\sqrt{m^2+\sigma^2} \\ 1 & \sigma \end{pmatrix} \begin{pmatrix} \tilde{w}_p \\ w_q \end{pmatrix}. \tag{E.11}$$

This completes the derivation of the resolvent: the four components of $z = (\sigma\mathbb{I} - \mathcal{L})^{-1} w$ are obtained from (4.37), (E.5) with (E.7), and (E.4). In particular,

$$\begin{aligned} u \;=\; & \frac{1}{D(\sigma)} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x|} (\sigma w_q + w_p) \\ & - \frac{\sigma^2+1}{2D(\sigma)\sqrt{m^2+\sigma^2}} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x|} \int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x'|} \left(\sigma w_u(x') + w_v(x')\right)\, dx' \\ & + \frac{1}{2\sqrt{m^2+\sigma^2}} \int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2+\sigma^2}|x-x'|} \left(\sigma w_u(x') + w_v(x')\right)\, dx'. \end{aligned} \tag{E.12}$$

E.2 Eigenfunctions and eigenvalues

Returning to the Cauchy formula used in (4.36), we can now write functions of $\mathcal{L}$ explicitly. We have

$$f(\mathcal{L})w = \frac{1}{2\mathrm{i}\pi} \oint f(\sigma) \left(\sigma\mathbb{I} - \mathcal{L}\right)^{-1} w\, d\sigma = \frac{1}{2\mathrm{i}\pi} \oint f(\sigma) z(\sigma)\, dz,$$

where $z(\sigma)$ is that derived in the previous section. We assume that $m^2 > 1$, so that both the discrete and continuous spectra contribute to the integral.

We compute the components of $f(\mathcal{L})w$ in succession. We start by the contribution of the pole $\mathrm{i}\omega_*$ to $(f(\mathcal{L})w)_p$ found using L'Hôpital's rule to be

$$(f(\mathcal{L})w)_p = \frac{\mathrm{i}\omega_* f(\mathrm{i}\omega_*)}{D'(\mathrm{i}\omega_*)}\left[\mathrm{i}\omega_* w_q + w_p + T\int_{-\infty}^{\infty} \mathrm{e}^{-\sqrt{m^2-\omega_*^2}|x'|}\left(\mathrm{i}\omega_* w_u(x') + w_v(x')\right)\,dx'\right].$$

The contribution of the complex-conjugate pole $-\mathrm{i}\omega_*$ is obtained by replacing ω_* by $-\omega_*$. Having covered the contribution of the residues, we note that the rest of the integral is obtained by integrating along either side of each of the branch cuts that are present due to the multivalued square-root term. Useful references for this type of procedure are found in [1, 65], and more extensive examples of contour deformation for uncovering Landau poles are given in several works by Chapman [10, 11] and also Chapman and Sororkin [12] . To compute the contribution of the branch cut $\mathrm{i}[m,\infty)$, we let $\sigma = \mathrm{i}s$, with $s \in [m,\infty)$; then $\sqrt{m^2+\sigma^2} = \pm\mathrm{i}\sqrt{s^2-m^2}$, with $+$ [$-$] on the right [left] side of the cut. Also, $D(\sigma) \to D(\mathrm{i}s) := 1 - s^2 + 2\mathrm{i}T\sqrt{s^2-m^2}$ on the right, and $D(\sigma) \to \bar{D}(\mathrm{i}s)$ on the left. We obtain

$$\begin{aligned}
(f(\mathcal{L})w)_p &= \frac{1}{2\pi}\int_m^{\infty} f(\mathrm{i}s)\,ds\left\{(\mathrm{i}sw_p - w_q)\left(\frac{1}{D(\mathrm{i}s)} - \frac{1}{\bar{D}(\mathrm{i}s)}\right)\right.\\
&\qquad -2T\mathrm{i}\sqrt{s^2-m^2}w_q\left(\frac{1}{D(\mathrm{i}s)} + \frac{1}{\bar{D}(\mathrm{i}s)}\right)\\
&\qquad \left. + \mathrm{i}sT\int_{-\infty}^{\infty}\left(\frac{\mathrm{e}^{-\mathrm{i}\sqrt{s^2-m^2}|x'|}}{D(\mathrm{i}s)} - \frac{\mathrm{e}^{\mathrm{i}\sqrt{s^2-m^2}|x'|}}{\bar{D}(\mathrm{i}s)}\right)\left(\mathrm{i}sw_u(x') + w_v(x')\right)\,dx'\right\}\\
&= \frac{T}{\pi}\int_m^{\infty}\frac{sf(\mathrm{i}s)\,ds}{|D(\mathrm{i}s)|^2}\left\{2\sqrt{s^2-m^2}(\mathrm{i}sw_q + w_p) + \int_{-\infty}^{\infty}\phi(x';s)\left(\mathrm{i}sw_u(x') + w_v(x')\right)\,dx'\right\},
\end{aligned}$$

where we have defined

$$\phi(x';s) = 2T\sqrt{s^2-m^2}\cos\left(\sqrt{s^2-m^2}|x'|\right) + (1-s^2)\sin\left(\sqrt{s^2-m^2}|x'|\right).$$

The contribution from the cut $\mathrm{i}(-\infty,-m]$ is obtained by taking $s \mapsto -s$. A similar computation gives the q-component $(f(\mathcal{L}))_q$ as

$$(f(\mathcal{L})w)_q = \frac{f(\mathrm{i}\omega_*)}{D'(\mathrm{i}\omega_*)}\left[\mathrm{i}\omega_* w_q + w_p + T\int_{-\infty}^{\infty}\mathrm{e}^{-\sqrt{m^2-\omega_*^2}|x'|}\left(\mathrm{i}\omega_* w_u(x') + w_v(x')\right)\,dx'\right]$$

for the discrete spectrum contribution, and

$$(f(\mathcal{L})w)_q = -\frac{\mathrm{i}T}{\pi}\int_m^{\infty}\frac{f(\mathrm{i}s)\,ds}{|D(\mathrm{i}s)|^2}\left\{2\sqrt{s^2-m^2}(\mathrm{i}sw_q + w_p) + \int_{-\infty}^{\infty}\phi(x';s)\left(\mathrm{i}sw_u(x') + w_v(x')\right)\,dx'\right\}$$

for the continuous spectrum.

We now turn to $(f(\mathcal{L})w)_u$. From (4.38), we find the contribution of the discrete spectrum

$$(f(\mathcal{L})w)_u = \frac{f(\mathrm{i}\omega_*)}{D'(\mathrm{i}\omega_*)}\left[\mathrm{i}\omega_* w_q + w_p + T\int_{-\infty}^{\infty}\mathrm{e}^{-\sqrt{m^2-\omega_*^2}|x'|}\left(\mathrm{i}\omega_* w_u(x') + w_v(x')\right)\,dx'\right]\mathrm{e}^{-\sqrt{m^2-\omega_*^2}|x|}.$$

The continuous spectrum is computed as follows:

$$
\begin{aligned}
\left(f(\mathcal{L})w\right)_u &= \frac{1}{2\pi}\int_m^\infty f(\mathrm{i}s)\,ds\left\{(\mathrm{i}sw_q+w_p)\left(\frac{\mathrm{e}^{-\mathrm{i}\sqrt{s^2-m^2}|x|}}{D(\mathrm{i}s)}-\frac{\mathrm{e}^{\mathrm{i}\sqrt{s^2-m^2}|x|}}{\bar{D}(\mathrm{i}s)}\right)\right.\\
&\quad+\frac{\mathrm{i}(1-s^2)}{2\sqrt{s^2-m^2}}\int_{-\infty}^\infty\left(\frac{\mathrm{e}^{-\mathrm{i}\sqrt{s^2-m^2}(|x|+|x'|)}}{D(\mathrm{i}s)}+\frac{\mathrm{e}^{\mathrm{i}\sqrt{s^2-m^2}(|x|+|x'|)}}{\bar{D}(\mathrm{i}s)}\right)(\mathrm{i}sw_u(x')+w_v(x'))\;dx'\\
&\quad\left.-\frac{\mathrm{i}}{2\sqrt{s^2-m^2}}\int_{-\infty}^\infty\left(\mathrm{e}^{-\mathrm{i}\sqrt{s^2-m^2}(|x-x'|)}+\mathrm{e}^{\mathrm{i}\sqrt{s^2-m^2}(|x-x'|)}\right)(\mathrm{i}sw_u(x')+w_v(x'))\;dx'\right\}\\
&= -\frac{\mathrm{i}}{\pi}\int_m^\infty\frac{f(\mathrm{i}s)}{|D(\mathrm{i}s)|^2}\phi(x;s)(\mathrm{i}sw_q+w_p)\,ds\\
&\quad+\frac{\mathrm{i}}{2\pi}\int_m^\infty\frac{f(\mathrm{i}s)}{\sqrt{s^2-m^2}|D(\mathrm{i}s)|^2}\,ds\int_{-\infty}^\infty(\mathrm{i}sw_u(x')+w_v(x'))\;dx'\\
&\quad\times\Big[(1-s^2)^2\cos\left(\sqrt{s^2-m^2}(|x|+|x'|)\right)\\
&\qquad-4T(1-s^2)\sqrt{s^2-m^2}\sin\left(\sqrt{s^2-m^2}(|x|+|x'|)\right)\\
&\qquad-\left((1-s^2)^2+4T^2(s^2-m^2)\right)\cos\left(\sqrt{s^2-m^2}|x-x'|\right)\Big]\\
&= -\frac{\mathrm{i}}{\pi}\int_m^\infty\frac{f(\mathrm{i}s)}{|D(\mathrm{i}s)|^2}\phi(x;s)(\mathrm{i}sw_q+w_p)\,ds\\
&\quad-\frac{\mathrm{i}}{2\pi}\int_m^\infty\frac{f(\mathrm{i}s)\phi(x;s)}{\sqrt{s^2-m^2}|D(\mathrm{i}s)|^2}\,ds\int_{-\infty}^\infty\phi(x';s)\,(\mathrm{i}sw_u(x')+w_v(x'))\;dx'\\
&\quad-\frac{\mathrm{i}}{2\pi}\int_m^\infty\frac{f(\mathrm{i}s)\psi(x;s)}{\sqrt{s^2-m^2}}\,ds\int_{-\infty}^\infty\psi(x';s)\,(\mathrm{i}sw_u(x')+w_v(x'))\;dx',
\end{aligned}
$$

where we have introduced

$$\psi(x;s)=\sin\left(\sqrt{s^2-m^2}\,x\right).$$

Similar expressions result for $\left(f(\mathcal{L}w)\right)_v$; they correspond to multiplication of $f(\mathrm{i}\omega_*)$ by $\mathrm{i}\omega_*$ and $f(\mathrm{i}s)$ by $\mathrm{i}s$.

Introducing the inner product, the results are of the form

$$\langle z_1,z_2\rangle=\tilde{p}_1p_2+\bar{q}_1q_2+\int_{-\infty}^\infty[\bar{v}_1(x)v_2(x)+\bar{u}_1(x)u_2(x)]dx. \tag{E.13}$$

The results above are of the form

$$f(\mathcal{L})w=\sum_k f(\sigma_k)\frac{\langle\tilde{e}_k,w\rangle}{\tilde{e}_k,e_k}e_k, \tag{E.14}$$

where the sum runs over the spectrum and should be interpreted as an integral for the continuous part. This makes it possible to identify the eigenvectors (or spectral projectors), together with their duals, and the normalisation factors. Specifically, the discrete eigenvalue $\mathrm{i}\omega$ has its eigenvector and its dual given by

$$
e_*=\begin{pmatrix}p\\q\\v(x)\\u(x)\end{pmatrix}=\begin{pmatrix}1\\-\mathrm{i}\omega_*^{-1}\\\mathrm{e}^{-\sqrt{m^2-\omega_*^2|x|}}\\-\mathrm{i}\omega_*^{-1}\mathrm{e}^{-\sqrt{m^2-\omega_*^2|x|}}\end{pmatrix}\quad\text{and}\quad\tilde{e}_*=\begin{pmatrix}\tilde{p}\\\tilde{q}\\\tilde{v}(x)\\\tilde{u}(x)\end{pmatrix}=\begin{pmatrix}1\\-\mathrm{i}\omega_*\\T\mathrm{e}^{-\sqrt{m^2-\omega_*^2|x|}}\\-\mathrm{i}\omega_*T\mathrm{e}^{-\sqrt{m^2-\omega_*^2|x|}}\end{pmatrix}. \tag{E.15}
$$

The continuous spectrum has two families of 'eigenfunctions': first

$$e_{1s} = \begin{pmatrix} 2T \\ -2\mathrm{i}s^{-1}T \\ \phi(x;s)/\sqrt{s^2-m^2} \\ -\mathrm{i}s\phi(x:s/\sqrt{s^2-m^2}) \end{pmatrix} \quad \text{and} \quad \tilde{e}_{1s} = \begin{pmatrix} 2T \\ -2\mathrm{i}s^{-1}T \\ \phi(x;s)/\sqrt{s^2-m^2} \\ -\mathrm{i}s\phi(x:s/\sqrt{s^2-m^2}) \end{pmatrix}, \tag{E.16}$$

and second

$$e_{2s} = \begin{pmatrix} 0 \\ 0 \\ T\psi(x;s)/\sqrt{s^2-m^2} \\ -\mathrm{i}s^{-1}T\psi(x:s/\sqrt{s^2-m^2}) \end{pmatrix} \quad \text{and} \quad \tilde{e}_{2s} = \begin{pmatrix} 0 \\ 0 \\ T\psi(x;s)/\sqrt{s^2-m^2} \\ -\mathrm{i}sT\psi(x:s/\sqrt{s^2-m^2}) \end{pmatrix}. \tag{E.17}$$

The two complex conjugates are also eigenfunctions. The normalisation factors can be read off as

$$\langle \tilde{e}_*, e_* \rangle = \frac{D'(\mathrm{i}\omega}{\mathrm{i}\omega_*}, \quad \langle \tilde{e}_{1s}, e_{1s'} \rangle = \frac{2\pi T|D(\mathrm{i}s)|^2}{s\sqrt{s^2-m^2}}\delta(s-s'), \quad \langle \tilde{e}_{2s}, e_{2s} \rangle = \frac{2\pi T\sqrt{s^2-m^2}}{s}\delta(s-s'). \tag{E.18}$$

All the other inner products of eigenvectors vanish.

E.3 Calculating the resolvent for the damped system

As before in order to find the resolvent we solve the following system for z, where

$$z(\sigma) = (p(\sigma), q(\sigma), u(x,\sigma), v(x,\sigma)) : \tag{E.19}$$

$$\sigma p + q - T[u_x]_-^+ = w_p, \tag{E.20}$$

$$\sigma q - p = w_q, \tag{E.21}$$

$$\sigma v - u_{xx} + m^2 u + nv = w_v, \tag{E.22}$$

$$\sigma u - v = w_u. \tag{E.23}$$

Begin by taking (E.22) + $(\sigma + n)\times$(E.23); this gives:

$$u_{xx} - (m^2 + \sigma(\sigma+n))u = -w_v - (\sigma+n)w_u. \tag{E.24}$$

The Green's function is

$$G(x,x') = -\frac{1}{2\sqrt{m^2+\sigma(\sigma+n)}} e^{-\sqrt{m^2+\sigma(\sigma+n)}|x-x'|}, \tag{E.25}$$

and so the general solution can be written as

$$u(x) = Ce^{-\sqrt{m^2+\sigma(\sigma+n)}|x|} + \frac{1}{2\sqrt{m^2+\sigma(\sigma+n)}} \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma(\sigma+n)}|x-x'|}[(\sigma+n)w_u + w_v]dx'. \tag{E.26}$$

Consider

$$m^2 + \sigma(\sigma+n) = 0; \tag{E.27}$$

solving for σ yields the branch points,

$$\sigma_B = -\frac{n}{2} \pm \frac{\sqrt{n^2 - 4m^2}}{2}. \tag{E.28}$$

Now $q(t) = u(x = 0, t)$, hence

$$q = C + \frac{1}{2\sqrt{m^2 + \sigma(\sigma + n)}} \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma(\sigma+n)}|x'|}[(\sigma + n)w_u + w_v]dx', \tag{E.29}$$

and together with (E.26) the solution obtained is

$$u(x) = qe^{-\sqrt{m^2+\sigma(\sigma+n)}|x|} + \frac{1}{2\sqrt{m^2 + \sigma(\sigma + n)}} \int_{-\infty}^{\infty} H(x, x'; \sigma)[(\sigma + n)w_u(x') + w_v(x')]dx', \tag{E.30}$$

where

$$H(x, x'; \sigma) = e^{-\sqrt{m^2+\sigma(\sigma+n)}|x-x'|} - e^{-\sqrt{m^2+\sigma(\sigma+n)}|x|+|x'|}. \tag{E.31}$$

Now using this in (E.30) gives

$$[u_x]^+_- = -2\sqrt{m^2 + \sigma(\sigma + n)}q + \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma(\sigma+n)}|x'|}[(\sigma + n)w_u + w_v]dx', \tag{E.32}$$

and together with (E.20) we obtain

$$\sigma p + q + 2T\sqrt{m^2 + \sigma(\sigma + n)}q = \tilde{w}_p, \tag{E.33}$$

where

$$\tilde{w}_p = w_p + T \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma(\sigma+n)}|x'|}[(\sigma + n)w_u + w_v]dx'. \tag{E.34}$$

Solving (E.21) and (E.33) for p and q gives

$$p = \frac{1}{1 + \sigma^2 + 2T\sqrt{m^2 + \sigma(\sigma + n)}} \left[\left(-1 - 2T\sqrt{m^2 + \sigma(\sigma + n)}\right) w_q + \sigma\tilde{w}_p\right] \tag{E.35}$$

and

$$q = \frac{1}{1 + \sigma^2 + 2T\sqrt{m^2 + \sigma(\sigma + n)}}(\sigma w_q + \tilde{w}_p). \tag{E.36}$$

For $n < 2m$ the branch points have a nonzero imaginary part, and we make cuts on $[-\frac{n}{2} + \sqrt{n^2 - 4m^2}, \infty)$ and $(-\infty, -\frac{n}{2} - \sqrt{n^2 - 4m^2}]$ (fig. 4.18). When integrating we parametrise the contour using s as before

$$\sigma = -n/2 + \mathrm{i}s; \quad s \in [\tfrac{1}{2}\sqrt{4m^2 - n^2}, \infty), \tag{E.37}$$

which changes the form of the dispersion relation

$$\begin{aligned} D(s) &= 1 + \sigma(s)^2 \pm 2T\sqrt{m^2 + \sigma(s)(\sigma(s) + n)} \\ &= 1 + n^2/4 - n\mathrm{i}s - s^2 \pm 2T\mathrm{i}\sqrt{s^2 + n^2/4 - m^2}. \end{aligned}$$

We are now able to compute the contributions to the spectrum from both the poles and the branch cuts. In what follows we assume $w_u = w_u(x')$ and $w_v = w_v(x')$ for brevity. For $f(\mathcal{L}(w))_p$ the contribution from the pole is calculated by using L'Hôpital's rule:

$$f(\mathcal{L}(w))_p = \frac{\sigma_* f(\sigma_*)}{D'(\sigma_*)} \left[w_p + \sigma_* w_q + T \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma_*(\sigma_*+n)}|x'|} [(\sigma_* + n)w_u + w_v] dx' \right]. \quad \text{(E.38)}$$

The contribution from the continuous spectrum is given by integrating around the branch cut:

$$f(\mathcal{L}(w))_p = \frac{T}{\pi} \int_{\frac{1}{2}\sqrt{4m^2-n^2}}^{\infty} \frac{f(-n/2+\mathrm{i}s)ds\ (-n/2+\mathrm{i}s)}{D\tilde{D}} \left\{ 2\sqrt{s^2+n^2/4-m^2}[w_p \right.$$
$$\left. +(-n/2+\mathrm{i}s)w_q] + \int_{-\infty}^{\infty} \phi(x';s)[(n/2+\mathrm{i}s)w_u + w_v]dx' \right\}. \quad \text{(E.39)}$$

Here D and $\tilde{D}$ denote $D(\sigma)$ on the right- and left-hand side of the cut, respectively; i.e.

$$D = 1 + n^2/4 - n\mathrm{i}s - s^2 + 2T\mathrm{i}\sqrt{s^2+n^2/4-m^2}, \quad \text{(E.40)}$$
$$\tilde{D} = 1 + n^2/4 - n\mathrm{i}s - s^2 - 2T\mathrm{i}\sqrt{s^2+n^2/4-m^2}. \quad \text{(E.41)}$$

Note that $D\tilde{D} \neq |D|^2$, in contrast to the case where $n = 0$. Also,

$$\phi(x';s) = 2T\sqrt{s^2+n^2/4-m^2}\cos\left(\sqrt{s^2+n^2/4-m^2}|x'|\right)$$
$$+ (1 + n^2/4 - n\mathrm{i}s - s^2)\sin\left(\sqrt{s^2+n^2/4-m^2}|x'|\right).$$

Similarly, the next pole contribution is

$$(f(\mathcal{L})w)_q = \frac{f(\sigma_*)}{D'(\sigma_*)} \left[w_p + \sigma_* w_q + T \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma_*(\sigma_*+n)}|x'|} [(\sigma_* + n)w_u + w_v] dx' \right], \quad \text{(E.42)}$$

and that from the continuous spectrum is

$$(f(\mathcal{L})w)_q = -\frac{\mathrm{i}T}{\pi} \int_{\frac{1}{2}\sqrt{4m^2-n^2}ds}^{\infty} \frac{f(-n/2+\mathrm{i}s)ds}{D\tilde{D}} \left\{ 2\sqrt{s^2+n^2/4-m^2}[w_p \right.$$
$$\left. +(-n/2+\mathrm{i}s)w_q] + \int_{-\infty}^{\infty} \phi(x';s)[(n/2+\mathrm{i}s)w_u + w_v]dx' \right\}. \quad \text{(E.43)}$$

Finally for the case of $(f(\mathcal{L})w)_u$, the part from the poles is given by

$$(f(\mathcal{L})w)_u = \frac{f(\sigma_*)}{D'(\sigma_*)} \Big[\sigma_* w_q + w_p$$
$$+ T \int_{-\infty}^{\infty} e^{-\sqrt{m^2+\sigma_*(\sigma_*+n)}|x'|} [(\sigma_* + n)w_u + w_v] dx' \Big]$$
$$\times e^{-\sqrt{m^2+\sigma_*(\sigma_*+n)}|x|}, \quad \text{(E.44)}$$

and the part from the branch cut is

$$
\begin{aligned}
(f(\mathcal{L})w)_q = \quad & - \quad \frac{\mathrm{i}}{\pi}\int_{\frac{1}{2}\sqrt{4m^2-n^2}}^{\infty} \frac{f(-n/2+\mathrm{i}s)ds}{\sqrt{s^2+n^2/4-m^2}D\tilde{D}}\phi(x;s)[(n/2+\mathrm{i}s)w_q+w_p]ds \qquad \text{(E.45)}\\
& - \quad \frac{i}{2\pi}\int_{\frac{1}{2}\sqrt{4m^2-n^2}}^{\infty} \frac{f(-n/2+\mathrm{i}s)ds\,\phi(x;s)}{\sqrt{s^2+n^2/4-m^2}D\tilde{D}}\int_{-\infty}^{\infty}\phi(x';s)[(n/2+\mathrm{i}s)w_u+w_v]dx'\\
& - \quad \frac{i}{2\pi}\int_{\frac{1}{2}\sqrt{4m^2-n^2}}^{\infty} \frac{f(-n/2+\mathrm{i}s)ds\,\phi(x;s)}{\sqrt{s^2+n^2/4-m^2}}\int_{-\infty}^{\infty}\psi(x';s)[(n/2+\mathrm{i}s)w_u+w_v]dx'.
\end{aligned}
$$

Here we have defined

$$\psi(x;s) = \sin(\sqrt{s^2+n^2/4-m^2}|x|). \qquad \text{(E.46)}$$

To find the contributions from the other branch, it is only necessary to replace s by $-s$ in the preceding calculations. A similar result to (E.45) can be derived for $(f(\mathcal{L})w)_v$.

Appendix F

Wave absorbing sponge layers for the shallow water equations

F.1 Motivation

In this thesis we have considered in detail two specific methods for simulating freely propagating waves, in one and two dimensions. The Klein-Gordon equation was handled using an exact discretely non-reflecting boundary condition, whereas the nonlinear shallow water equations made use of an approximate higher order boundary condition. The latter — involving a Higdon NRBC applied to the linear part of the equations — relied on the nonlinear activity being confined to a "core" of activity, maintaining a reasonable separation from the boundary, and hence allowing the boundary condition to neglect the presence of nonlinear terms. For some problems this assumption may not hold, and another approach must be taken. One such approach, as mentioned earlier, is to simply enlarge the computational domain enough that even the fastest waves supported by the differencing scheme cannot reach the boundary and risk reflection. However, this quickly becomes infeasible for simulations of any great duration. Instead, in this section we will consider the introduction of *wave absorbing layers*, more commonly termed *sponge layers*.

Sponge layers are widely employed in fluid solvers (see for example, [45, 58]). Their properties and effectiveness have been studied numerically by many, including Lavelle and Thacker [51] and more recently by Modave et al [60]. A sponge layer is a region of damping a few (usually ~ 10) grid points in thickness. The damping takes the form of Rayleigh friction, and is implemented as an extra term in each of the equations for $\partial_t u$, $\partial_t v$ and $\partial_t \eta$. These terms are respectively, $-\sigma_x(x)u$, $-\sigma_y(y)v$ and $(-\sigma_x(x) + \sigma_y(y))\eta$, where $\sigma_x(x)$, $\sigma_y(y)$ are damping constants that depend on the position of the grid point (on each of the staggered grids) at which the friction is evaluated. In this way, the damping of a particular sponge has a specific profile, which will be seen to affect its effectiveness. The use of separate sponges σ_x and σ_y ensures that only velocity components **normal to a particular boundary** are damped. At the boundary of the model a simple zero Dirichlet boundary condition $u = v = \eta = 0$ is imposed, as it is hoped that any resulting reflected waves will have very little energy after having passed through the absorbing layer, and will have still less after the return journey.

F.2 The choice of sponge

Numerical tests show that the performance of such layers is affected by both their thickness δ_{sponge}, the level of damping, and sponge profile. Consider the case of $\sigma(x, y) = \sigma_{\text{constant}}$: with σ_{constant} too small the absorption will be ineffective; with a choice that is too large the interface between the areas of no damping and strong damping will act like a hard reflecting wall. As a result, in order to provide adequate attenuation of waves, it is wise to use a sponge with a small damping constant, spread over a larger sponge layer (of say, at least ten grid points). Experiments show that the best behaviour is obtained using a damping coefficient that varies gradually from zero at the interface to its maximum value at the boundary; this is sensible as the inner edge of the sponge layer no longer appears like a hard wall. Having chosen a layer width of a certain number of grid points, and a maximum value of damping σ_{m}, all that remains is to choose the functional form of $\sigma(x, y)$. In [60] several profiles were compared: the uniform profile $\sigma(x) = \sigma_{\text{m}} = \text{constant}$, the polynomial profile

$$\sigma(x) = \sigma_m \left(\frac{x}{\delta_{\text{sponge}}} \right)^{\alpha}, \tag{F.1}$$

where $\alpha > 0$ and σ_{m} is a constant, and also the shifted hyperbolic profile

$$\sigma(x) = \frac{\sqrt{gH}}{\delta_{\text{sponge}}} \frac{x}{\delta_{\text{sponge}} - x}. \tag{F.2}$$

We compare the same candidate profiles with various choices of σ_{m} for use in a fully nonlinear scheme, by means of the same experiment used in [60] but with a different difference scheme, in order to choose the best available sponge layer for a given width. The parameter σ_m can be tuned for each sponge layer (except the shifted hyperbolic case), in order to find the optimal sponge configuration. We first of all look at the experimental set-up considered in §4 of [60]. An initial perturbation height field was allowed to adjust without forcing, and the initial velocity fields were both zero. The f-plane approximation was used with $f = 1.024 \times 10^{-4}$, $g = 9.81\text{ms}^{-1}$, a resting height of $H = 100\text{m}$, $\Delta t = 50\text{s}$, and $\Delta x = \Delta y = 10^4\text{m}$. The inner domain was a rectangle of 110 points by 51 points, and the initial height field was of the form

$$\eta_0 = \exp\left(-\frac{x^2 + y^2}{2.5 \times 10^9} \right). \tag{F.3}$$

The wavespeed for this problem is given by $\sqrt{gH} = 31.3\text{ms}^{-1}$, and using this we estimate that some reflected waves should be returning to the centre of the domain after around 240 time steps, which we choose for the run length of the experiment.

In order to measure the reflection and absorption properties of the various sponges, we use the following definition

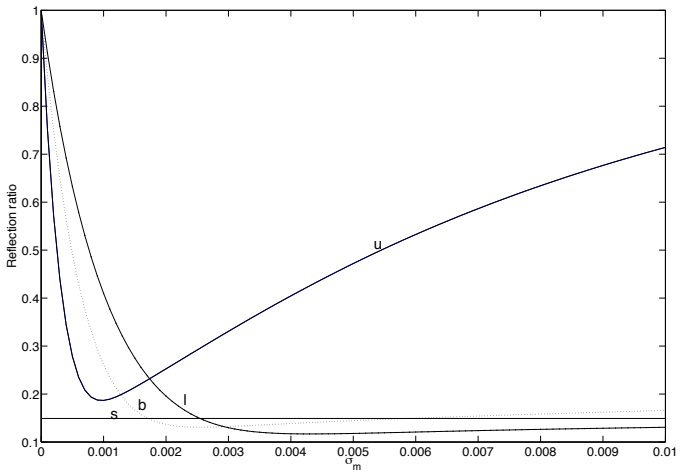
$$\text{Reflection ratio} = \sqrt{\frac{\text{Error at final time with sponge layer damping}}{\text{Error at final time with no damping in sponge layer}}}, \tag{F.4}$$

where

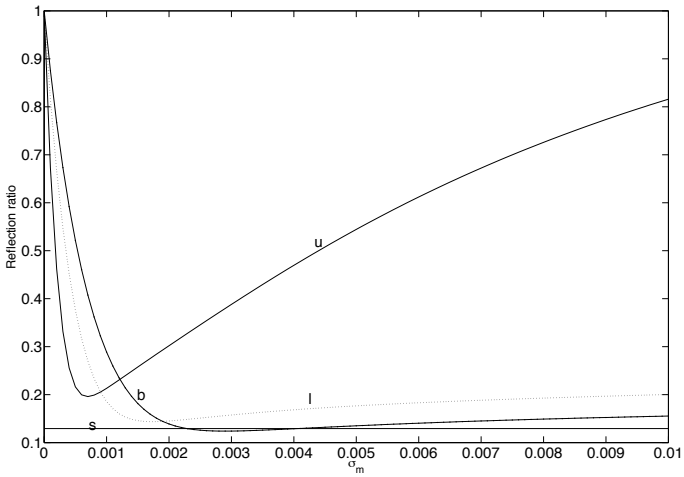
$$\text{Error}(t) = \int_\Omega \left(\frac{g}{2}[\eta(t) - \eta_{\text{ref}}(t)]^2 + \frac{h}{2}[u(t) - u_{\text{ref}}]^2 + \frac{h}{2}[v(t) - v_{\text{ref}}(t)]^2 \right) d\Omega, \tag{F.5}$$

and Ω is the domain excluding the sponge layers. The fields with the subscript "ref" are obtained by selecting part of the solution computed using a very large computational domain, i.e. they correspond to the exact discrete solution without reflection, since no waves reach the boundary during the history of the experiment. Despite requiring a large domain, these fields need only be computed once. The results differ slightly from those in [60], due to their use of the FBTCS time-scheme as opposed to second order Adams-Bashforth which is used for this work, a different interior spatial differencing scheme, and the choices for Δt, g and the number of time steps being different. Also in this work, the solution with a hard wall is computed by imposing the zero Dirichlet condition at the outer edge of the sponge layer, rather than the inner edge as in [60]. This makes is less meaningful to compare different thicknesses of sponge layer, as the hard wall solution will change as the sponge thickness is changed. However, we only wish to evaluate different profiles; it is well known that a thicker sponge should fare better.

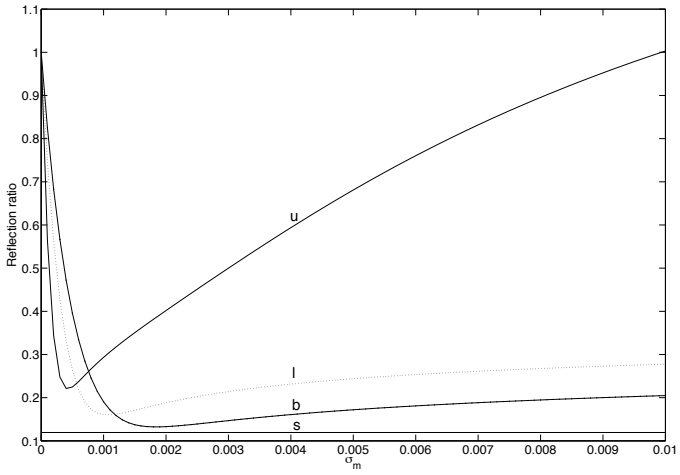
Figure F.1 shows plots of the reflection ratio for each of the profiles (uniform, linear, quadratic and shifted hyperbolic) as the maximum damping constant σ_{m} is varied. There are plots for sponges of five, seven, and ten grid points in thickness. Since the shifted hyperbolic profile has no free parameter, its associated reflection ratio is shown with a flat line. For each other profile, there is an optimal value of σ_{m} for which the reflection ratio is minimised. Notice that as the sponge layer is thickened, the shifted hyperbolic profile eventually performs best. The uniform profile is, as expected, always the worst performing profile.



(a) 5 point sponge.



(b) 7 point sponge.



(c) 10 point sponge.

Figure F.1: Plots of reflection ratios for sponge layers of various thickness, using several different profiles: uniform (u), linear(l), binomial (b), and shifted hyperbolic (s).

Computer code

The MATLAB and Fortran90 source code for the programs and functions used to generate most of the numerical results in this thesis are to be found on the accompanying compact disc. In most cases, the only documentation that should be required can be found throughout the source files in the form of comments. Compilation and basic use instructions for the Fortran90 shallow water code are given in the file `README.txt`. The various codes used in each chapter are briefly described below.

The shallow water code `shallowf.f90` is an OpenMP Fortran90 program based on software freely available from the National Centre for Atmospheric Research, written by Paul Swarztrauber and subsequently modified by Juliana Rew. The `fftpack` suite was written by Paul Swarztrauber at NCAR, and the `random_normal` function was the work of Alan Miller. All other code is the author's.

Chapter 2

1-d problem

- `recurrence.m`
  Generates coefficients for the exact DNRBC using a recurrence relation.
- `nrbc_coeff.m`
  Generates coefficients for the exact DNRBC using the FFT, to specified accuracy.
- `kappa_optimize.m`, `ref_fun.m`
  Calling program used to generate optimal coefficients for the fixed-length NRBC, and helper function which returns a measure of the discrete reflection ratio.
- `discrete_ref.m`
  Solves the Klein-Gordon equation using a specific DNRBC, and quantifies its accuracy.
- `nrbc_coeff_asselin.m`
  Generates coefficients for the exact DNRBC using the FFT, taking account of Asselin filtering.

2-d problem

- `shallowStaggeredNrbc.m`
  Staggered-time/staggered-space linearised shallow water solver featuring Higdon higher-order boundary conditions.

- `shallowMatNrbcFun.m`
  Function to generate the update matrix as outlined in chapter 2 and Appendix A.
- `higbc.m`
  Function to implement algorithm 2 of chapter 2.
- `givoli.m`
  Function to determine the free parameters c_j in the Higdon boundary condition.
- `grid2vec.m`, `hmat.m`, `huv2vec.m`, `num2arr.m`, `num2tri.m`, `uvmat.m`, `vec2huv.m`
  Helper functions required by `shallowMatNrbcFun.m` and `shallowStaggeredNrbc.m`.
- `pseudotrace.m`
  Kostin-Brühl pseudospectral boundary tracing algorithm.

Chapter 3

- `extlk1_flux_runs.m`
  Numerical solver for the forced eLK model.
- `gaussf.m`
  Generates smooth Gaussian random forcing with given autocorrelation; helper function for `extlk1_flux_runs.m`
- `welchper.m`
  Welch periodogram function.
- `shallowf.f90`
  2-d nonlinear rotating shallow water solver, as outlined in Appendix C (see `README.txt`).
- `swef.m`
  MATLAB wrapper code for simulating forced shallow water with the Fortran90 shallow water program; generates initial fields, calls `shallowf` and reads and processes results.

Chapter 4

- `lamb_energy_filtered.m`
  Lamb model solver. Computes solution and theoretical decay rate estimates for comparison.

Appendix B

- `qvh_schecter.m`
  Function to determine basic state fields for an axisymmetric vortex given a radial PV distribution.
- `qvh_error.m`
  Helper function for `qvh_schecter.m` that implements the shooting method.
- `disturbance_schecter.m`
  Linear solver for 1-d vortex time-evolution problem in shallow water.

Bibliography

[1] M. J. Ablowitz and A. S. Fokas, *Complex variables: Introduction and applications*, Cambridge Univ Pr, 2003.

[2] X. Antoine, C. Besse, and P. Klein, *Absorbing boundary conditions for the one-dimensional schrödinger equation with an exterior repulsive potential*, J. Comput. Phys., 228 (2009), pp. 312–335.

[3] R. Asselin, *Frequency filter for time integrations*, Mon. Weather. Rev., 100 (1972), pp. 487–490.

[4] N. J. Balmforth, S. G. Llewellyn Smith, and W. R. Young, *Disturbing vortices*, J. Fluid Mech., 426 (2001), pp. 95–133.

[5] A. M. Bloch, P. Hagerty, A. G. Rojo, and M. I. Weinstein, *Gyroscopically stabilized oscillators and heat baths*, J. Stat. Phys., 115 (2004), pp. 1073–1100.

[6] R. J. Briggs, J. D. Daugherty, and R. H. Levy, *Role of Landau damping in crossed-field electron beams and inviscid shear flow*, Phys. Fluids, 13 (1970), p. 421.

[7] E. G. Broadbent and D. W. Moore, *Acoustic destabilization of vortices*, Philos. T. Roy. Soc. A, 290 (1979), pp. 353–371.

[8] G. Browning, H. O. Kreiss, and J. Oliger, *Mesh refinement*, Math. Comput., 27 (1973), pp. 29–39.

[9] M. Brühl, *A curve tracing algorithm for computing the pseudospectrum*, BIT Numer. Math., 36 (1996), pp. 441–454.

[10] C. J. Chapman, *High-speed leading-edge noise*, P. Roy. Soc. Lond. A Mat, 459 (2003), p. 2131.

[11] ——, *Some benchmark problems for computational aeroacoustics*, J. Sound Vib., 270 (2004), pp. 495–508.

[12] C. J. Chapman and S. V. Sorokin, *The forced vibration of an elastic plate under significant fluid loading*, J. Sound Vib., 281 (2005), pp. 719–741.

[13] E. Danioux, J. Vanneste, P. Klein, and H. Sasaki, *Spontaneous inertia-gravity-wave generation by surface-intensified turbulence*, Journal of Fluid Mechanics, 699 (2012), p. 153.

[14] J. R. Dea, *High-order non-reflecting boundary conditions for the linearized Euler equations*, PhD thesis, Naval Postgraduate School, Montery, CA, 2008.

[15] ——, *An experimental adaptation of higdon-type non-reflecting boundary conditions to linear first-order systems*, J. Comput. Appl. Math., 235 (2011), pp. 1354–1366.

[16] J. R. Dea, F. X. Giraldo, and B. Neta, *High-order non-reflecting boundary conditions for the linearized 2-d euler equations: No mean flow case*, Wave Motion, 46 (2009), pp. 210–220.

[17] D. B. Duncan, *Symplectic finite difference approximations of the nonlinear klein-gordon equation*, SIAM J. Numer. Anal., 34 (1997), pp. pp. 1742–1760.

[18] M. Durand and F. El Dabaghi, *High performance computing II*, North-Holland, 1991.

[19] D. R. Durran, *Numerical methods for wave equations in geophysical fluid dynamics*, vol. 32, Springer Verlag, 1999.

[20] M. Ehrhardt, *Discrete artificial boundary conditions*, PhD thesis, TU Berlin, 2001.

[21] ——, *Absorbing boundary conditions for hyperbolic systems*, Numer. Math. Theor. Meth. Appl., 3 (2010), pp. 295–337.

[22] A. Eliassen, *The quasi-static equations of motion with pressure as independent variable*, Grøndahl & sons boktr., I kommisjon hos Cammermeyers boghandel, 1949.

[23] B. Engquist and A. Majda, *Radiation boundary conditions for acoustic and elastic wave calculations*, Comm. Pure Appl. Math., 32 (1979), pp. 313–357.

[24] R. Ford, *The instability of an axisymmetric vortex with monotonic potential vorticity in rotating shallow water*, J. Fluid Mech., 280 (1994), pp. 303–334.

[25] R. Ford, M. E. McIntyre, and W. A. Norton, *Balance and the slow quasimanifold: Some explicit results*, J. Atmos. Sci., 57 (2000), pp. 1236–1254.

[26] A. E. Gill, *Atmosphere-ocean dynamics*, International Geophysics Series, Academic Press, London, 1982.

[27] D. Givoli, *Non-reflecting boundary conditions*, J. Comput. Phys., 94 (1991), pp. 1 – 29.

[28] ——, *High-order local non-reflecting boundary conditions: A review*, Wave Motion, 39 (2004), pp. 319 – 326. New computational methods for wave propagation.

[29] D. Givoli and B. Neta, *High-order non-reflecting boundary conditions for dispersive waves*, Wave Motion, 37 (2003), pp. 257–271.

[30] ——, *High-order nonreflecting boundary conditions for the dispersive shallow water equations*, J. Comput. Appl. Math., 158 (2003), pp. 49–60.

[31] M. J. Grote and J. B. Keller, *Exact nonreflecting boundary conditions for the time dependent wave equation*, SIAM J. Appl. Math., (1995), pp. 280–297.

[32] B. Gustafsson, H. O. Kreiss, and A. Sundström, *Stability theory of difference approximations for mixed initial boundary value problems. II*, Math. Comput., 26 (1972), pp. 649–686.

[33] P. Hagerty, A. M. Bloch, and M. I. Weinstein, *Radiation induced instability*, SIAM J. App. Math., 64 (2003), pp. 484–524.

[34] T. Hagstrom, *On high-order radiation boundary conditions*, in Computational Wave Propagation, B. Engquist and G. Kriegsmann, eds., The IMA Volumes in Mathematics and its Applications, Springer-Verlag, New York, 1997, pp. 1–22.

[35] ——, *Radiation boundary conditions for the numerical simulation of waves*, Acta Numer., 8 (1999), pp. 47–106.

[36] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration*, Springer, 2005.

[37] L. Halpern, *Absorbing boundary conditions for the discretization schemes of the one-dimensional wave equation*, Math. Comput., 38 (1982), pp. 415–429.

[38] H. Han and Z. Zhang, *An analysis of the finite-difference method for one-dimensional Klein-Gordon equation on unbounded domain*, Appl. Numer. Math., 59 (2009), pp. 1568–1583.

[39] R. L. Higdon, *Absorbing boundary conditions for difference approximations to the multi-dimensional wave equation*, Math. Comput., 47 (1986), pp. 437–459.

[40] ——, *Numerical absorbing boundary conditions for the wave equation*, Math. Comput., 49 (1987), pp. 65–90.

[41] ——, *Radiation boundary conditions for elastic wave propagation*, SIAM J. Numer. Anal., (1990), pp. 831–869.

[42] ——, *Absorbing boundary conditions for acoustic and elastic waves in stratified media*, J. Comput. Phys., 101 (1992), pp. 386–418.

[43] ——, *Radiation boundary conditions for dispersive waves*, SIAM J. Numer. Anal., 31 (1994), pp. 64–100.

[44] J. R. Holton, ed., *Encyclopedia of atmospheric sciences*, Academic Press, Oxford, 2003.

[45] M. Israeli and S. A. Orszag, *Approximation of radiation boundary conditions*, J. Comput. Phys., 41 (1981), pp. 115–135.

[46] E. I. Jury, *Theory and application of the z-transform method*, Wiley, 1964.

[47] S. K. Kar and R. P. Turco, *Formulation of a lateral sponge layer for limited-area shallow-water models and an extension for the vertically stratified case*, Mon. Weather. Rev., 123 (1995), pp. 1542–1559.

[48] J. B. Keller and D. Givoli, *Exact non-reflecting boundary conditions*, J. Comput. Phys., 82 (1989), pp. 172 – 192.

[49] V. F. Kop'ev and E. A. Leonte'ev, *Acoustic instability of an axial vortex*, Sov. Phys. Acoust, 29 (1983), pp. 111–115.

[50] H. Lamb, *On a peculiarity of the wave-system due to the free vibrations of a nucleus in an extended medium*, P. Lond. Math. Soc., 1 (1900), p. 208.

[51] J. W. Lavelle and W. C. Thacker, *A pretty good sponge: Dealing with open boundaries in limited-area ocean models*, Ocean Mod., 20 (2008), pp. 270–292.

[52] C. E. Leith, *Nonlinear normal mode initialization and quasi-geostrophic theory*, J. Atmos. Sci., 37 (1980), pp. 958–968.

[53] G. Lill, *Diskrete Randbedingungen an knstlichen Rndern*, PhD thesis, TU Darmstadt, 1992.

[54] E. N. Lorenz, *Attractor sets and quasi-geostrophic equilibrium.*, J. Atmos. Sci., 37 (1980), pp. 1685–1699.

[55] ——, *On the existence of a slow manifold*, J. Atmos. Sci., 43 (1986), pp. 1547–1557.

[56] E. N. Lorenz and V. Krishnamurthy, *On the nonexistence of a slow manifold*, J. Atmos. Sci., 44 (1987), pp. 2940–2950.

[57] S. H. Lui, *Computation of pseudospectra by continuation*, SIAM J. Sci. Comput., 18 (1997), pp. 565–573.

[58] P. Marchesiello, J. C. McWilliams, and A. Shchepetkin, *Open boundary conditions for long-term integration of regional oceanic models*, Ocean Mod., 3 (2001), pp. 1–20.

[59] M. E. McIntyre, *Global effects of gravity waves in the middle atmosphere: A theoretical perspective*, Advances in Space Research, 27 (2001), pp. 1723–1736.

[60] A. Modave, É. Deleersnijder, and É. J. M. Delhez, *On the parameters of absorbing layers for shallow water models*, Ocean Dyn., 60 (2010), pp. 65–79.

[61] B. Neta, V. van Joolen, J. R. Dea, and D. Givoli, *Application of high-order higdon non-reflecting boundary conditions to linear shallow water models*, Commun. Numer. Meth. En., 24 (2008), pp. 1459–1466.

[62] J. Nycander and K. Döös, *Open boundary conditions for barotropic waves*, J. Geophys. Res., 108 (2003).

[63] H. Nyquist, *Regeneration theory*, AT&T Tech. J., 11 (1932), pp. 126–147.

[64] G. M. Reznik, V. Zeitlin, and M. Ben Jelloul, *Nonlinear theory of geostrophic adjustment. part 1. rotating shallow-water model*, J. Fluid Mech., 445 (2001), pp. 93–120.

[65] K. F. Riley, M. P. Hobson, and S. J. Bence, *Mathematical methods for physics and engineering*, Cambridge Univ Pr, 2002.

[66] A. J. Robert, *The integration of a low order spectral form of the primitive meteorological equations*, J. Meteor. Soc. Japan, 44 (1966), pp. 237–245.

[67] C. W. Rowley and T. Colonius, *Discretely Nonreflecting Boundary Conditions for Linear Hyperbolic Systems*, J. Comput. Phys., 157 (2000), pp. 500–538.

[68] R. Sadourny, *The dynamics of finite-difference models of the shallow-water equations*, J. Atmos. Sci., (1975), pp. 680–689.

[69] D. A. Schecter and M. T. Montgomery, *Conditions that inhibit the spontaneous radiation of spiral inertia-gravity waves from an intense mesoscale cyclone*, J. Atmos. Sci., 63 (2006), pp. 435–456.

[70] A. Schuster, *On the investigation of hidden periodicities with application to a supposed twenty-six-day period of meterological phenomena.*, Terr. Mag., 3 (1898), pp. 13–41.

[71] J. Thuburn, *Numerical wave propagation on the hexagonal c-grid*, J. Comput. Phys., 227 (2008), pp. 5836–5858.

[72] L. N. Trefethen, *Group velocity interpretation of the stability theory of Gustafsson, Kreiss, and Sundström*, J. Comput. Phys., 49 (1983), pp. 199–217.

[73] L. N. Trefethen, *Computation of pseudospectra*, Acta numerica, 8 (1999), pp. 247–295.

[74] L. N. Trefethen and M. Embree, *Spectra and pseudospectra: The behavior of nonnormal matrices and operators*, Princeton Univ Pr, 2005.

[75] S. V. Tsynkov, *Numerical solution of problems on unbounded domains. a review*, Appl. Numer. Math., 27 (1998), pp. 465–532.

[76] J. Tuomela and O. Vacus, *Algebraic derivation of discrete absorbing boundary conditions for the wave equation*, Numer. Math., 80 (1998), pp. 419–459.

[77] M. R. Turner and A. D. Gilbert, *Linear and nonlinear decay of cat's eyes in two-dimensional vortices, and the link to Landau poles*, J. Fluid Mech., 593 (2007), pp. 255–279.

[78] V. J. Van Joolen, *Application of Higdon non-reflecting boundary conditions to shallow water models*, PhD thesis, Naval Postgraduate School Monterey CA, 2003.

[79] J. Vanneste, *Wave radiation by balanced motion in a simple model*, SIAM J. Appl. Dyn. Syst, 5 (2006), pp. 783–807.

[80] ——, *Exponential smallness of inertia-gravity wave generation at small rossby number*, J. Atmos. Sci., 65 (2008), pp. 1622–1637.

[81] J. Vanneste, *Balance and spontaneous wave generation in geophysical flows*, Annual Review of Fluid Mechanics, 45 (2012).

[82] J. Vanneste and I. Yavneh, *Exponentially small inertia-gravity waves and the breakdown of quasigeostrophic balance*, Journal of the atmospheric sciences, 61 (2004), pp. 211–223.

[83] L. Wagatha, *On boundary conditions for the numerical simulation of wave propagation*, Applied Numerical Mathematics, 1 (1985), pp. 309–314.

[84] T. Warn, O. Bokhove, T. G. Shepherd, and G. K. Vallis, *Rossby number expansions, slaving principles, and balance dynamics*, Q. J. Roy. Meteor. Soc., 121 (1995), pp. 723–739.

[85] P. Welch, *The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms*, IEEE T. Aud. Elec., 15 (1967), pp. 70–73.

[86] P. D. Williams, T. W. Haine, and P. L. Read, *Inertia-gravity waves emitted from balanced flow: observations, properties, and consequences*, Journal of the Atmospheric Sciences, 65 (2008), pp. 3543–3556.

[87] J. C. Wilson, *Derivation of boundary conditions for the artificial boundaries associated with the solution of certain time dependent problems by lax–wendroff type difference schemes*, P. Edinburgh Math. Soc., 25 (1982), pp. 1–18.

[88] T. G. Wright, *Eigtool*, Software available at http://www. comlab. ox. ac. uk/pseudospectra/eigtool, (2002).

[89] D. W. Zingg, *Aspects of linear stability analysis for higher-order finite-difference methods*, order, 4 (1997), p. 6.

[90] D. W. Zingg and M. Lederle, *On linear stability analysis of high-order finite-difference methods*, AIAA paper, 5249 (2005), p. 2005.

[91] D. W. Zingg, H. Lomax, and H. Jurgens, *High-accuracy finite-difference schemes for linear wave propagation*, SIAM J. Sci. Comput., 17 (1996), p. 328.